



RECEIVED
SEP 18 2001
Technology Center 2600

DECLARATION

I, MIYOKO KODAMA, residing at 23-5, Tateno-cho, Nerima-ku, Tokyo, Japan, do hereby solemnly and sincerely declare:

1. that I am well acquainted with the Japanese and English languages, and
2. that the attached document, translation of Copy of the Specification is a true and correct translation into the English language.

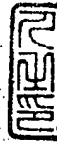
And I make this solemn declaration conscientiously believing the same to be true and correct.

Dated this 1st day of September, 2001.

RECEIVED
SEP 21 2001
Technology Center 2100

Miyoko Kodama

By: Miyoko Kodama



APOSTILLE

(Convention de La Haye du 5 octobre 1961)

1. Country: JAPAN

This public document

2. has been signed by Toshiaki HOUKIN

3. acting in the capacity of Director of the Tokyo Legal Affairs Bureau

4. bears the seal/stamp of

Certified

5. at Tokyo

6.

7th September, 2001.

7. by the Ministry of Foreign Affairs

8. 01 - No 022389

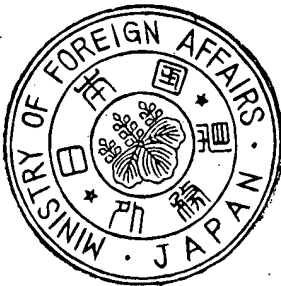
9. Seal/stamp:

10. Signature:

T. Mochizuki

Takeshi MOCHIZUKI

For the Minister for Foreign Affairs



平成 13 年 登 簿 第 2117 号

認 証

嘱 託 人 児玉美代子（住所 東京都練馬区立野町 2 3 - 5）は、
添付証書の署名が自己のものに相違ない旨、代理人 佐藤鈴子 を通じ本公証人
に対し自認した。

よつて、これを認証する。

平成 13 年 9 月 7 日、本公証人役場において
東京都千代田区麹町 5 丁目 2 番地 1

東京法務局所属
公証人

中重正人

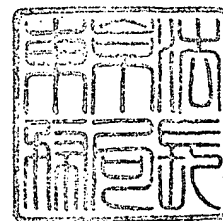
証 明

上記署名は、東京法務局所属公証人の署名に相違ないものであ
り、かつ、その押印は、真実のものであることを証明する。

平成 13 年 9 月 7 日

東京法務局長

寶金敏明

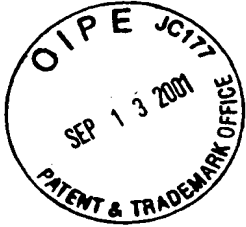
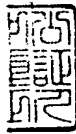


CERTIFICATE

This is to certify that the signature affixed above has been provided
by Notary, duly authorized by the Tokyo Legal Affairs Bureau and that
the Official Seal appearing on the same is genuine.

Date 7th September, 2001.

Toshiaki HOUKIN
Director of the Tokyo Legal Affairs Bureau



NOTARIAL CERTIFICATE

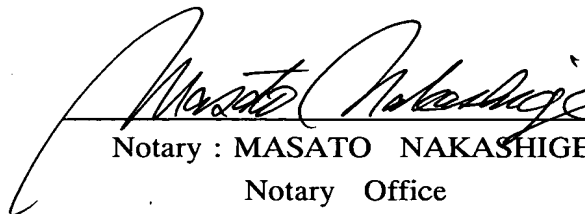
RECEIVED

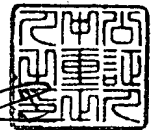
SEP 21 2001

Technology Center 2100

This is to certify that Ms.SUZUKO SATO, an agent of
Ms.MIYOKO KODAMA, residing at 23-5, Tateno-cho, Nerima-ku, Tokyo, Japan,
has stated that said Ms.MIYOKO KODAMA, acknowledged herself that the
signature on the attached document is her own.

Dated this 7th day of September, 2001.


Notary : MASATO NAKASHIGE



Notary Office
5-2-1 Kojimachi, Chiyoda-ku, Tokyo
Tokyo Legal Affairs Bureau



【書類名】 特許願

【整理番号】 9701389

【提出日】 平成9年9月8日

【あて先】 特許庁長官殿

【国際特許分類】 G06F 17/00

【発明の名称】 関連文書表示装置

【請求項の数】 10

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 内野 寛治

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 津田 宏

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 松井 くにお

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100074099

【郵便番号】 102

【住所又は居所】 東京都千代田区二番町8番地20 二番町ビル3F

【弁理士】

【氏名又は名称】 大菅 義之

【電話番号】 03-3238-0031

【選任した代理人】

【識別番号】 1 0 0 0 6 7 9 8 7

【郵便番号】 2 2 2

【住所又は居所】 神奈川県横浜市港北区太尾町 1 4 1 8 - 3 0 5 (大倉山二番館)

【弁理士】

【氏名又は名称】 久木元 彰

【電話番号】 0 4 5 - 5 4 5 - 9 2 8 0

【手数料の表示】

【予納台帳番号】 0 1 2 5 4 2

【納付金額】 2 1 0 0 0

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9 7 0 5 0 4 7

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 関連文書表示装置

【特許請求の範囲】

【請求項1】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

前記文書群情報を参照することにより、前記文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する文書群キーワード抽出手段と、

前記文書群毎に、それに関連するタイトルと、該文書群に対して前記文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する文書群キーワード表示手段と、

を含むことを特徴とする関連文書表示装置。

【請求項2】 請求項1に記載の装置であって、

前記文書群キーワード表示手段は、ユーザが指定した検索指定に該当する文書が含まれる文書群又は該検索指定に該当するキーワードを強調表示する、

ことを特徴とする関連文書表示装置。

【請求項3】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

前記文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報を抽出する文書属性解析手段と、

前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する文書群構造表示手段と、

を含むことを特徴とする関連文書表示装置。

【請求項 4】 請求項 3 に記載の装置であって、

前記文書群構造表示手段は、前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各省略された作者名と各文書量情報とが組み合せられて各ノードとして表示されるツリー構造によって表示する、

ことを特徴とする関連文書表示装置。

【請求項 5】 請求項 3 に記載の装置であって、

前記文書群構造表示手段は、前記文書群毎に、その文書群に含まれる文書から抽出される複数のトピックを、その文書群に対応して表示される前記ツリー構造を構成する前記各ノードと関連付けて、更に表示する、

ことを特徴とする関連文書表示装置。

【請求項 6】 請求項 5 に記載の装置であって、

前記文書群構造表示手段は、前記文書群毎に、前記各トピックとそれに関連する前記各ノードとを、同一の色で表示する、

ことを特徴とする関連文書表示装置。

【請求項 7】 請求項 3 に記載の装置であって、

前記文書群構造表示手段は、ユーザが指定した検索指定に該当する文書に対応するノードを強調表示する、

ことを特徴とする関連文書表示装置。

【請求項 8】 参照関係を有する文書からなる文書群を表示する関連文書表示装置であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する 1 つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する文書群解析手段と、

該分類された各文書群を、それぞれを構成する文書から抽出されるトピックに基づいて更に細分類し、その細分類に関する情報をトピック分類情報として抽出するトピック解析手段と、

前記文書群情報及び前記トピック分類情報を参照することにより、前記文書群

毎及びそれを細分類した前記トピック毎に、そのトピックに関連する文書に含まれるキーワードを抽出するトピックキーワード抽出手段と、

前記文書群毎及びそれを細分類した前記トピック毎に、それに関連するタイトルと、該トピックに対して前記トピックキーワード抽出手段によって抽出されたキーワードとを対応付けて表示するトピックキーワード表示手段と、
を含むことを特徴とする関連文書表示装置。

【請求項 9】 コンピュータにより使用されたときにそれによって読み出されるプログラムを記録した記録媒体であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する 1 つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する機能と、

前記文書群情報を参照することにより、前記文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する機能と、

前記文書群毎に、それに関連するタイトルと、該文書群に対して前記文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する機能と

を前記コンピュータに行わせるためのプログラムを記録したコンピュータ読出し可能記録媒体。

【請求項 10】 コンピュータにより使用されたときにそれによって読み出されるプログラムを記録した記録媒体であって、

文書集合を構成する複数の文書間の参照関係を解析することによって、前記文書集合を構成する複数の文書を参照関係を有する 1 つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する機能と、

前記文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報を抽出する機能と、

前記文書群情報及び前記文書属性情報を参照することによって、前記文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する機能と、

を前記コンピュータに行わせるためのプログラムを記録したコンピュータ読出し可能記録媒体。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、参照関係にある文書群を整理し、その文書群をいろいろな観点から表示することによって、ユーザによる必要な情報へのアクセスを支援する技術に関する。

【 0 0 0 2 】

【従来の技術及び発明が解決しようとする課題】

パソコン通信又はコンピュータネットワーク上で運用される電子会議室や電子ニュース等において、順次蓄積される文書集合の中から必要な文書をより迅速かつ簡単に見つけ出したいという要請が、従来からある。

【 0 0 0 3 】

このような要請に対して、文書集合中の各文書のタイトルを作成日順に並び替え、その結果得られるタイトルリストをユーザに掲示するという従来技術が知られている。

【 0 0 0 4 】

また、文書集合を互いに参照関係のある文書から構成される文書群に分類し、文書群中の各文書のタイトルをインデントして表示することにより、各文書の参照関係を掲示する従来技術や、文書群中の各文書の番号をツリーで表示することにより各文書の参照関係を掲示する従来技術も知られている。

【 0 0 0 5 】

更に、文書集合の中から、特定のキーワードを含む文書を全文検索し、その検索結果を羅列的に掲示する従来技術も知られている。

しかし、これらの従来技術におけるような限定された情報表示のみでは、以下に示される問題点を解決することができなかった。

1. 雑多な文書集合の中から必要な文書又は必要な文書が含まれているであろう文書群を見つけ出すためには、掲示される文書のタイトルに頼るしか

い。タイトルは、必ずしも文書内容を正確に表わしているとは限らないため、正確な検索が困難である。

2. タイトルのインデント表示や文書番号のツリー表示だけでは、文書群全体の構造を把握すること、及び文書群における話題の推移を把握することが困難である。
3. 種々の観点から必要な文書にアクセスすることができない。
4. 検索結果が多数件ある場合に、更に絞り込み検索を実行するか検索結果のリストを1件1件チェックしなければ、必要な文書にアクセスすることができない。

一方、複数の特定の文書からキーワードを抽出し、共通のキーワードを含む各文書に対し自動的に他の文書へのリンクを設定する技術が、知られている。この従来技術は、特許の公知例や研究論文等の特定文書中で互いの文献を相互参照することを可能にすることによって、関連する複数の文書を効率的に読み広げることが可能にする。

【 0 0 0 6 】

しかし、このような従来技術は、関連する文書の参照を容易にすることを目的としており、電子会議室や電子ニュース等の文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索の支援に、適用することはできなかった。

【 0 0 0 7 】

本発明の課題は、大量の文書が含まれる文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索を、支援することにある。

【 0 0 0 8 】

【課題を解決するための手段】

本発明は、参照関係を有する文書からなる文書群を表示する関連文書表示装置を前提とする。

【 0 0 0 9 】

本発明の第1の態様は、以下の構成を有する。

まず、文書群解析手段（構造解析部402）は、文書集合を構成する複数の文

書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群（スレッド）に分類し、その分類に関する情報を文書群情報（スレッドインデックス 3 0 5）として抽出する。

【 0 0 1 0 】

文書群キーワード抽出手段（内容解析部 4 0 3）は、文書群情報を参照することにより、文書群毎に、その文書群を構成する文書に含まれるキーワードを抽出する。

【 0 0 1 1 】

文書群キーワード表示手段（表示装置 3 0 6）は、文書群毎に、それに関連するタイトルと、その文書群に対して文書群キーワード抽出手段によって抽出されたキーワードとを対応付けて表示する。

【 0 0 1 2 】

上述した本発明の第 1 の態様の構成により、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろう文書群を容易に見つけ出すことが可能となる。

【 0 0 1 3 】

本発明の第 2 の態様は、上述の本発明の第 1 の態様の構成において、文書群キーワード表示手段が、ユーザが指定した検索指定に該当する文書が含まれる文書群又はその検索指定に該当するキーワードを強調表示する構成を有する。

【 0 0 1 4 】

上述した本発明の第 2 の態様の構成により、検索指定に該当する文書群を一目で把握することが可能となる。

本発明の第 3 の態様は、以下の構成を有する。

【 0 0 1 5 】

文書群解析手段（構造解析部 4 0 2）は、文書集合を構成する複数の文書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する1つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する。

【 0 0 1 6 】

文書属性解析手段（書式解析部 4 0 1）は、文書集合を構成する複数の文書のそれぞれから、その文書の属性を示す文書属性情報（メタインデックス 3 0 4）を抽出する。

【 0 0 1 7 】

文書群構造表示手段（表示装置 3 0 6）は、文書群情報及び文書属性情報を参照することによって、文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各文書属性情報又はその省略された情報が各ノードとして表示されるツリー構造によって表示する。より具体的には、文書群構造表示手段は、文書群毎に、その文書群の参照関係を、その文書群を構成する各文書に対応する各省略された作者名と各文書量情報（各行数）とが組み合せられて各ノードとして表示されるツリー構造によって表示する。

【 0 0 1 8 】

上述した本発明の第 3 の態様の構成により、文書群全体の構造を容易に把握することが可能となる。

本発明の第 4 の態様は、上述の本発明の第 3 の態様の構成において、文書群構造表示手段が、文書群毎に、その文書群に含まれる文書から抽出される複数のトピック（メイントピック、サブトピック）を、その文書群に対応して表示されるツリー構造を構成する各ノードと関連付けて、更に表示する構成を有する。より具体的には、文書群構造表示手段は、文書群毎に、各トピックとそれに関連する各ノードとを、同一の色で表示する。

【 0 0 1 9 】

上述した本発明の第 4 の態様の構成により、文書群全体の話題の推移を容易に把握することが可能となる。

本発明の第 5 の態様は、前述の本発明の第 3 の態様の構成において、文書群構造表示手段が、ユーザが指定した検索指定に該当する文書に対応するノードを強調表示する構成を有する。

【 0 0 2 0 】

上述した本発明の第 5 の態様の構成により、スレッドの構造を把握しつつ、検索を実行することが可能となる。

本発明の第 6 の態様は、以下の構成を有する。

【 0 0 2 1 】

文書群解析手段（構造解析部 4 0 2）は、文書集合を構成する複数の文書間の参照関係を解析することによって、文書集合を構成する複数の文書を参照関係を有する 1 つ以上の文書群に分類し、その分類に関する情報を文書群情報として抽出する。

【 0 0 2 2 】

トピック解析手段（構造解析部 4 0 2）は、その分類された各文書群を、それぞれを構成する文書から抽出されるトピックに基づいて更に細分類し、その細分類に関する情報をトピック分類情報（サブトピックインデックス）として抽出する。

【 0 0 2 3 】

トピックキーワード抽出手段（内容解析部 4 0 3）は、文書群情報及びトピック分類情報を参照することにより、文書群毎及びそれを細分類したトピック毎に、そのトピックに関連する文書に含まれるキーワードを抽出する。

【 0 0 2 4 】

トピックキーワード表示手段（表示装置 3 0 6）は、文書群毎及びそれを細分類したトピック毎に、それに関連するタイトルと、そのトピックに対してトピックキーワード抽出手段により抽出されたキーワードとを対応付けて表示する。

【 0 0 2 5 】

上述した本発明の第 6 の態様の構成により、キーワードによるより精密なトピックの絞込みを行うことが可能となる。

なお、本発明は、コンピュータにより使用されたときに、上述の本発明の各構成によって実現される機能と同様の機能をコンピュータに行わせるためのコンピュータ読出し可能記録媒体として構成することもできる。

【 0 0 2 6 】

【発明の実施の形態】

以下、図面を参照しながら本発明の実施の形態について詳細に説明する。

全体構成

図1は、本発明の実施の形態が対象とする、文書集合及び文書群の例を示す図である。

【 0 0 2 7 】

この例では、コンピュータネットワーク上に、ユーザが話題別に議論を行うフォーラムと呼ばれる仮想的な公開討論会場が設けられており、各フォーラムは、会議室と呼ばれる、更に細分化された話題を扱う複数の仮想的な会場に分類されている。ユーザがこの会議室に、発言文書を投稿（アップロード）することによって、議論が進行する。フォーラム及び会議室とも、サーバコンピュータ上のストレージエリアとして構成され、それらには、上述した分類基準に従って、文書が蓄積される。また、各会議室において、互いに参照関係を有する複数の文書からなる文書群がスレッドを構成する。

【 0 0 2 8 】

ユーザが投稿する文書は、例えば、図2に示されるデータ構造を有しており、その文書の番号を示す文書番号、日付、タイトル、その文書が参照する文書の番号である参照文書番号、作者名（発言者名）等の文書の属性フィールドが記載されるヘッダ部と、文書の本体が記載される内容部とから構成されている。

【 0 0 2 9 】

本発明の実施の形態では、以下のような表示形態が可能となる。

1. キーワードビュー：会議室を構成する文書集合において、その文書集合中の各スレッド毎に、そのスレッドを構成する文書群からキーワードが抽出され、それらのキーワードが、それらの文書数及びそれらが含まれるスレッドのタイトルと合わせて、図25に示される表示形態で表示される。

【 0 0 3 0 】

キーワードビューによって、ユーザは、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろうスレッド（文書群）を容易に見つけ出すことが可能となる。

2. スレッドビュー：文書の参照関係、タイトル、作者名、及び行数が一目にわかる図26に示される表示形態で、各スレッドを構成する文書群が表示される。

【 0 0 3 1 】

スレッドビューにより、スレッド全体の構造を把握し話題の推移を容易に把握することが可能となる。

3. 発言者ビュー：各文書のタイトルが、発言者（作者）毎に分類され、かつ発言者が発言の多い順にソートされ、同一発言者内では日付順で、図 2 7 に示される表示形態で、表示される。

【 0 0 3 2 】

発言者及び発言日付という観点から、文書集合（会議室）内の文書を参照することが可能となる。

4. 各ビューへの検索結果の反映：ユーザが指定した検索キーワードに関連する文書が、図 3 2 又は図 3 3 に示される表示形態で、キーワードビュー、スレッドビュー等の表示中で強調表示される。

【 0 0 3 3 】

この結果、より正確な文書の把握が可能となる。

5. 各ビューの切替え機能：上述のキーワードビュー、スレッドビュー、及び発言者ビューが任意に切替え可能とされることにより、種々の観点から必要な文書にアクセス可能となる。

以上のような表示形態を可能とする本発明の実施の形態について、詳細に説明する。

【 0 0 3 4 】

図 3 及び図 4 は、本発明の実施の形態のシステム構成図である。

フォーラム／会議室内の文書群は、所定のサーバコンピュータ上の文書群データベース 3 0 1 として、蓄積される。

【 0 0 3 5 】

文書群解析装置 3 0 2 は、文書群データベース 3 0 1 内の各会議室に対応する文書集合毎に、それに含まれる文書群の解析を行う。

集計装置 3 0 3 は、文書群解析装置 3 0 2 による解析結果に基づいて、メタインデックス 3 0 4、スレッドインデックス 3 0 5、及び索引ファイル 4 0 4 を生成する。

【0036】

表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【0037】

また、文字列検索装置405は、ユーザによる検索語の指定に基づいて、索引ファイル404を用いながら文書群データベース301内の文書集合を構成する各文書に対して検索を実行する。表示装置306は、その検索結果を、キーワードビュー又はスレッドビューに反映させて表示する。

【0038】

表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【0039】

文書群解析装置302は、書式解析部401、構造解析部402、及び内容解析部403とから構成される。

書式解析部401は、文書群データベース301内の文書集合を構成する図2のデータ構造を有する各文書のヘッダ部から、文書番号、タイトル、作者名、日付、及び参照文書番号を抽出し、また、各文書の内容部の行数を算出し、それらを、集計装置303を経由して、図5に示されるデータ構造を有するメタインデックス304に登録する。

【0040】

構造解析部402は、書式解析部401が各文書から抽出した文書番号と参照文書番号に基づいて、各文書をスレッドを単位とする文書群に分類し、集計装置303を経由して、スレッド毎に、それを構成する文書の参照関係のリストであるスレッドインデックス305を作成する。

【0041】

図6は、スレッドインデックス305のデータ構造を示す図である。

各スレッド毎に、ルート文書番号と、文書数と、スレッドの構造を示すリスト

とが登録される。リストは、

(親文書番号 子文書番号／サブツリー 子文書番号／サブツリー....)
という記述形式によって記述され、” 子文書番号／サブツリー” の部分には、更に再帰的 (リカーシブ) に、子リストを記述することができる。

【 0 0 4 2 】

図 6 に例示される 2 つのスレッドの各リストにより表現される各参照関係は、同図の表の右側に示される如くである。

また、構造解析部 4 0 2 は、解析したスレッドを構成する文書群を、更にタイトルを同一とするサブ文書群に分類し、各サブ文書群に色番号を付与し、各サブ文書群に含まれる文書に対応する図 5 のデータ構造を有するメタインデックス 3 0 4 内のエントリに、その文書が属するサブ文書群に付与された色番号を登録する。

【 0 0 4 3 】

内容解析部 4 0 3 は、構造解析部 4 0 2 によって分類されたスレッド毎に、そのスレッドを構成する文書群を 1 つの結合文書ファイルにまとめ、その結合文書からキーワードを抽出する。日本語文書からキーワードを抽出する技術としては、種々の公知技術を採用することができる。この場合に、キーワード抽出の精度を向上させるために、ノイズとなる文字がパターンマッチングによって除去される。また、例えば、上位所定個数のキーワードのみが抽出される。

【 0 0 4 4 】

内容解析部 4 0 3 によって抽出された各スレッドのキーワードは、集計装置 3 0 3 を経由して、そのスレッドのルート文書に対応する図 5 のデータ構造を有するメタインデックス 3 0 4 のエントリに、登録される。

【 0 0 4 5 】

また、内容解析部 4 0 3 は、スレッド毎に抽出したキーワードから、そのキーワードに含まれる索引語を抽出し、集計装置 3 0 3 を経由して、図 7 に示されるデータ構造を有する索引ファイル 4 0 4 を生成する。

【 0 0 4 6 】

この索引ファイル 4 0 4 は、前述したように、文字列検索装置 4 0 5 によって

参照される。

文書群解析装置302の詳細説明

図8は、図4の文書群解析装置302内の書式解析部401及び構造解析部402が実現する制御動作を示す動作フローチャートである。

【0047】

まず、書式解析部401は、文書群データベース301から新たに登録された新規文書ファイルから文書データを1行ずつ読み込みながら、その文書ファイルのヘッダ部（図2参照）から、文書番号、タイトル、作者名、日付、及び参照文書番号を抽出する（ステップ801→802→803→801のループ）。

【0048】

書式解析部401は、ヘッダ部の抽出を終了すると、集計装置303を経由して、図5のデータ構造を有するメタインデックス304において新規エントリを生成し、そのエントリに、抽出した文書番号、タイトル、作者名、日付、及び参照文書番号を登録する（ステップ802→804）。

【0049】

次に、書式解析部401は、上記新規文書ファイル内のヘッダ部以降の内容部（図2参照）から文書データを1行ずつ読み込みながら、文書末尾（EOF：エンドオブファイル）が検出されるまで、内容部の行数をカウントする（ステップ805→806→807→805のループ）。

【0050】

書式解析部401は、文書末尾を検出すると、集計装置303を経由して、それまでにカウントした内容部の行数を、図5のデータ構造を有するメタインデックス304の、現在処理中の新規文書の文書番号に対応するエントリに、登録する（ステップ806→808）。

【0051】

なお、上述の行数のカウント処理において、他の文書から引用している行（例えば”>”で始まる行）については、行数のカウントには算入しないことによって、その文書が実質的に発言している行数をカウントするように構成されてもよい。

【 0 0 5 2 】

続いて、構造解析部 4 0 2 に制御が移り、構造解析部 4 0 2 は、まず、現在処理中の新規文書の文書番号を、集計装置 3 0 3 を経由して、図 6 のデータ構造を有するスレッドインデックス 3 0 5 に登録する（ステップ 8 0 9）。

【 0 0 5 3 】

図 9 は、上記ステップ 8 0 9 の登録動作を示す動作フローチャートである。

まず、構造解析部 4 0 2 は、現在処理中の新規文書が、或るスレッドのルート文書であるか否かを判定する（ステップ 9 0 1）。具体的には、構造解析部 4 0 2 は、図 8 のステップ 8 0 1～8 0 3 のループにおいて、現在処理中の新規文書から参照文書番号が検出されなかった場合に、その文書はルート文書であると判定する。

【 0 0 5 4 】

構造解析部 4 0 2 は、現在処理中の新規文書が或るスレッドのルート文書であると判定した場合には、集計装置 3 0 3 を経由して、スレッドインデックス 3 0 5 において新規エントリを生成し、そのエントリに現在処理中の新規文書の文書番号をルート文書番号として登録する（ステップ 9 0 1→9 0 2）。

【 0 0 5 5 】

構造解析部 4 0 2 は、ステップ 9 0 2 の処理の後、上記エントリの文書数を 1 に初期設定し（ステップ 9 0 6）、図 8 のステップ 8 0 9 の処理を終了する。

一方、構造解析部 4 0 2 は、現在処理中の新規文書が或るスレッドのルート文書ではないと判定した場合には、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号が含まれるスレッドインデックス 3 0 5 中のエントリに、その参照文書番号を親文書番号とするリストが存在するか否かを判定する（ステップ 9 0 1→9 0 3）。

【 0 0 5 6 】

構造解析部 4 0 2 は、上述のエントリに、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号を親文書番号とするリストが存在すると判定した場合には、現在処理中の新規文書のヘッダ部から抽出されている文書番号を、そのリストの子文書番号として登録する（ステップ 9 0 3→9 0 5）。

【0057】

一方、構造解析部402は、上述のエントリに、現在処理中の新規文書のヘッダ部から抽出されている参照文書番号を親文書番号とするリストが存在しないと判定した場合には、そのエントリに、その参照文書番号を親文書番号とするリストを生成した上で、現在処理中の新規文書のヘッダ部から抽出されている文書番号を、そのリストの子文書番号として登録する（ステップ903→904→905）。

【0058】

構造解析部402は、ステップ905の処理の後、上述のエントリの文書数を更新（プラス1）し（ステップ906）、図8のステップ809の処理を終了する。

【0059】

上記図9の動作フローチャートによって実現される制御動作の具体例につき、図10の説明図を用いて説明する。この図は、図6のスレッドインデックス305において、ルート文書番号が“001”であるスレッドのエントリのリストが生成される過程を示すものである。

【0060】

まず、文書番号“001”のルート文書が処理される時点で、図9のステップ901→902が実行されることにより、スレッドインデックス305において新規エントリが生成され、そのエントリに文書番号“001”がルート文書番号として登録され（図10の（1））、上記エントリの文書数が1に初期設定される。

【0061】

次に、文書番号“002”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号“001”のエントリにおいて、文書番号“002”の文書から抽出された参照文書番号“001”を親文書番号とするリストが生成された後（図10の（2））、文書番号“002”がそのリストの子文書番号として登録され（図10の（3）の下線部）、上記エントリの文書数が2に変更

新される。

【0062】

次に、文書番号”003”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”003”の文書から抽出された参照文書番号”002”を親文書番号とするリストが生成された後（図10の（4）の下線部）、文書番号”003”がそのリストの子文書番号として登録され（図10の（5）の下線部）、上記エントリの文書数が3に更新される。

【0063】

次に、文書番号”004”の文書が処理される時点で、図9のステップ901→903→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”004”の文書から抽出された参照文書番号”001”を親文書番号とするリストの子文書番号として、文書番号”004”が登録され（図10の（6）の下線部）、上記エントリの文書数が4に更新される。

【0064】

次に、文書番号”005”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”005”の文書から抽出された参照文書番号”004”を親文書番号とするリストが生成された後（図10の（7）の下線部）、文書番号”005”がそのリストの子文書番号として登録され（図10の（8）の下線部）、上記エントリの文書数が5に更新される。

【0065】

最後に、文書番号”006”の文書が処理される時点で、図9のステップ901→903→904→905→906が実行されることにより、スレッドインデックス305内のルート文書番号”001”のエントリにおいて、文書番号”006”の文書から抽出された参照文書番号”005”を親文書番号とするリスト

が生成された後(図10の(9)の下線部)、文書番号"006"がそのリストの子文書番号として登録され(図10の(10)の下線部)、上記エントリの文書数が6に更新される。

【0066】

以上説明した図8のステップ809の処理の後、構造解析部402は、現在処理中の新規文書について色番号を決定し、その色番号を図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに登録する処理を実行する(図8のステップ810)。

【0067】

図11は上記ステップ810の登録動作を示す動作フローチャートである。なお、この登録動作では、図12に示されるデータ構造を有するカラーテーブルが使用される。このテーブルは、特には図示しない記憶装置に記憶される。

【0068】

まず、構造解析部402は、現在処理中の新規文書が、或るスレッドのルート文書であるか否かを判定する(ステップ1101)。具体的には、構造解析部402は、図8のステップ801~803のループにおいて、現在処理中の新規文書から参照文書番号が検出されなかった場合に、その文書はルート文書であると判定する。

【0069】

構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書であると判定した場合は、そのルート文書の文書番号に対応するエントリを、図12のデータ構造を有するカラーテーブルに登録し、そのエントリに現在処理中の新規文書から抽出された文書番号及びタイトル(図8のステップ804参照)と、初期色番号を登録する(ステップ1101→1102→1103)。図12の例では、ルート文書番号"001"の色番号"#1"に対応するエントリが登録され、そのタイトルはメインピックとなり、また、そのエントリの文書番号フィールドには、当初はルート文書番号"001"のみが登録される。

【0070】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス3

04 中の上記新規文書の文書番号に対応するエントリに、ステップ1103で登録した初期色番号を登録し（ステップ1103→1110）、図8のステップ810の処理を終了する。

【0071】

一方、構造解析部402は、現在処理中の新規文書が或るスレッドのルート文書ではないと判定した場合には、現在処理中の新規文書から抽出されたタイトル（図8のステップ804参照）が、“Re:”等の参照記号を含んでいるか否かを判定する（ステップ1101→1104）。

【0072】

構造解析部402は、現在処理中の新規文書から抽出されたタイトルが参照記号を含んでいると判定した場合はそのタイトルから参照記号を削除し（ステップ1104→1105）、現在処理中の新規文書から抽出されたタイトルが参照記号を含んでいないと判定した場合にはステップ1105は実行しない。

【0073】

その後、構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れかのエントリに、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されているか否かを判定する（ステップ1106）。現在処理中の新規文書が属するスレッドとそのルート文書番号は、図8のステップ809の処理において図6のデータ構造を有するスレッドインデックス305のエントリが決定される際に検出されるため、そのルート文書番号からカラーテーブル中のエントリが決定される。例えば、現在処理中の新規文書が文書番号“002”の文書である場合には、図12に示されるカラーテーブルにおいて、ルート文書番号“001”に属するエントリが検出される。

【0074】

構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れかのエントリに、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されていると判定した場合には、そのエントリの文書番号フィールドに、現在処理中の

新規文書の文書番号を登録する（ステップ1106→1107）。例えば、現在処理中の新規文書が文書番号”002”の文書である場合には、図12に示されるカラーテーブルにおいて、ルート文書番号”001”に属し色番号”#1”が登録されているエントリの文書番号フィールドに、文書番号”002”が登録される。

【0075】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに、ステップ1107で登録が行われたカラーテーブル中のエントリに設定されている色番号を登録し（ステップ1107→1110）、図8のステップ810の処理を終了する。

【0076】

一方、構造解析部402は、図12のデータ構造を有するカラーテーブル中の現在処理中の新規文書が属するスレッドに対応する何れのエントリにも、現在処理中の新規文書から抽出され参照記号を含まないタイトルと同じタイトルが登録されてはいないと判定した場合は、カラーテーブルにおいて上記スレッドに対応する新たなエントリを作成し（ステップ1108）、その作成したエントリに、そのスレッド内で新たな色番号と、現在処理中の新規文書から抽出された文書番号及びタイトル（図8のステップ804参照）を登録する（ステップ1106→1108→1109）。例えば、現在処理中の新規文書が文書番号”003”の文書である場合には、図12のカラーテーブルにおいて、ルート文書番号”001”に属する新たなエントリが作成され、そのエントリに、色番号”#2”と、文書番号”003”の文書のタイトルと、文書番号”003”とが登録される。このタイトルは、ルート文書番号”001”のタイトルであるメインピックに対して、サブピック1となる。

【0077】

その後、構造解析部402は、図5のデータ構造を有するメタインデックス304中の上記新規文書の文書番号に対応するエントリに、ステップ1109でカラーテーブル中の新たなエントリに設定された新たな色番号を登録し（ステップ1109→1110）、図8のステップ810の処理を終了する。

【 0 0 7 8 】

内容解析部 4 0 3 は、図 6 のデータ構造を有するスレッドインデックス 3 0 5 を参照することにより、前述したように、スレッド毎に、そのスレッドを構成する文書群を 1 つの結合文書ファイルにまとめ、その結合文書からキーワードを抽出する。この結果、抽出された各スレッドのキーワードは、そのスレッドのルート文書に対応する図 5 のデータ構造を有するメタインデックス 3 0 4 のエントリに、登録される。

表示装置 3 0 6 の詳細説明

表示装置 3 0 6 は、前述したように、図 5 のデータ構造を有するメタインデックス 3 0 4 と図 6 のデータ構造を有するスレッドインデックス 3 0 5 を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示することができる。

【 0 0 7 9 】

ここで例えば、図 4 のシステムが、ホームページの表示を制御する W e b サーバに接続されるように構成されれば、ユーザは、パーソナルコンピュータ等の手元の端末上の W e b ブラウザアプリケーションから上記 W e b サーバに接続して特定のフォーラムの特定の会議室にログインした後に、所定の各 G U I (グラフィックユーザインタフェース) ボタンをマウス装置等でクリックすることによって、キーワードビュー、スレッドビュー、又は発言者ビューを切り替えて表示させることができる。

【 0 0 8 0 】

より具体的には、表示装置 3 0 6 は、W e b サーバに対して例えば C G I (コモンゲートウェイインタフェース) アプリケーションとして機能し、W e b サーバから引き渡されたユーザからのリクエストに应答して、キーワードビュー、スレッドビュー、又は発言者ビュー等の各ビューを表現する H T M L (ハイパーテキストマークアップ言語) による文書データを生成し、それを W e b サーバに引き渡す。そして、これらの H T M L 文書データを W e b サーバがユーザにインターネット等のコンピュータネットワークを経由して返信することにより、ユーザの端末上の W e b ブラウザアプリケーションに、上記ビューが表示される。

【0081】

まず、表示装置306が実現するキーワードビューの表示動作について説明する。

前述したようにキーワードビューにおいては、スレッド毎に、そのスレッドを構成する文書群から抽出されているキーワードが、その文書群の文書数及びそのスレッドのタイトルと合わせて、図25に示される表示形態で表示される。

【0082】

図13は、表示装置306が実行するキーワードビューの表示動作を示す動作フローチャートである。。

まず、表示装置306は、図5のデータ構造を有するメタインデックス304のファイルを読み込む（ステップ1301）。

【0083】

次に、表示装置306は、メタインデックス304のファイルから1エントリずつデータを読み込みながら、ルート文書が登録されているエントリを検索する（ステップ1301→1302→1301のループ）。各エントリがルート文書が登録されているエントリであるか否かは、各エントリの参照文書番号フィールドの値が無効なデータ値であるか否かによって判定することができる。

【0084】

表示装置306は、ルート文書が登録されているエントリを検出すると、そのルート文書番号を、そのルート文書番号に対応する文書群データベース301内のルート文書を表示するためのアプリケーションへの統一されたアドレス情報であるURL (Uniform Resource Locator) がHREF属性の値として指定されるアンカータグに変換する（ステップ1302→1303）。

【0085】

次に表示装置306は、図6のデータ構造を有するスレッドインデックス305において、上記ルート文書番号に対応するエントリを参照することにより、そのスレッドに含まれる文書数（子文書数）を取得する（ステップ1304）。

【0086】

そして、表示装置306は、図5のデータ構造を有するメタインデックス30

4において、上記ルート文書が登録されているエントリから、タイトル（メイントピック）と、キーワードとを抽出し、それらと、ステップ1303で変換されたアンカータグ形式のルート文書番号、及びステップ1304で取得した子文書数からなるデータ列を1テーブルレコードとして含むHTMLテーブル文書データを作成する（ステップ1305）。

【0087】

続いて、表示装置306は、メタインデックス304のファイルから文書末尾（EOF）を検出するまで、上記ステップ1301～1305の一連の処理を繰り返し実行することにより、各スレッド毎のHTMLテーブル文書データを作成する（ステップ1306→1301）。

【0088】

表示装置306は、メタインデックス304のファイルから文書末尾を検出すると（ステップ1306の判定がYES）、最終的に得られたHTMLテーブル文書データをWebサーバに引き渡して、キーワードビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図25に例示されるようなテーブル形式で、キーワードビューが表示される。

【0089】

ユーザは、キーワードビュー上の各スレッド毎のキーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろうスレッドを容易に見つけ出すことが可能となる。

【0090】

また、ユーザは、ルート文書に対応するアンカーをマウス装置等でクリックすることによって、所望のスレッドのルート文書に即座にアクセスすることができる。

【0091】

上述のキーワードビューの表示動作において、子文書数に応じて、各スレッドのテーブルレコードを色分けして表示するように構成されてもよい。これによって、ユーザは、スレッド毎の発言数を一目で判別することができる。

【0092】

続いて、表示装置 3 0 6 が実現するスレッドビューの表示動作について説明する。

前述したように、スレッドビューにおいては、文書の参照関係、タイトル、作者名、及び行数が一目にわかる図 2 6 に示される表示形態で、各スレッドを構成する文書群が表示される。

【 0 0 9 3 】

図 2 6 において、スレッドの参照関係及び話題の推移が色付きツリーによって表示される。各ツリーのノードは、各文書に対応し、その文書の作者名の先頭文字 (2 バイト) とその文書の行数を用いて、

作者名 [行数]

の形式で表示される。また、各ノードの前後には、" * " 、" + " 、" = " 、又は " . " 等の記号が付される。これらの記号の意味は、下記の通りである。

" * " この記号が付される文書がルート文書である。

【 0 0 9 4 】

" + " この記号が付される文書が参照している文書が他の文書によっても参照されている。

" = " この記号が付される文書を参照している文書が存在する。

【 0 0 9 5 】

" . " この記号が付される文書を参照している文書が存在しない。

また、図 2 6 において、" Main Topic : " に続いてそのスレッドのルート文書のタイトルが表示され、" Sub Topic : " に続いてそのスレッド中に現れるルート文書のタイトル以外のタイトルが表示される。そして、各タイトルは色分けされ、各タイトルと同じタイトル (参照記号を除く) を有する文書に対応するノードは、そのタイトルの色と同じ色で表示される。

【 0 0 9 6 】

これによって、ユーザは、スレッド全体の構造を把握しスレッド内の話題の推移を一目で把握することが可能となる。

更に、各ノードはアンカーとして表示される。これにより、ユーザは、各ノードをマウス装置等によってクリックすることにより、そのノードに対応する文書

に即座にアクセスすることができる。

【0097】

図14は、表示装置306が実行するスレッドビューの表示動作を示す動作フローチャートである。。

まず、表示装置306は、図6のデータ構造を有するスレッドインデックス305のファイルから、1つのスレッドに対応する1つのエントリ（1行）のリストと、そのスレッドに含まれる文書数を、読み込む（ステップ1401）。例えば、図6のデータ構造を有するスレッドインデックス305において、ルート文書番号”001”に対応するリストとして、

(001 (002 003) (004 (005 006)))

が読み込まれ、文書数として”6”が読み込まれる。

【0098】

次に、表示装置306は、読み込んだリストから、例えば図6の表の右側に示されるスレッドのツリー構造を復元する（ステップ1402）。このツリー構造を表現するために、表示装置306は、例えば図15に示されるような配列データを生成する。

【0099】

次に、表示装置306は、読み込んだリストの各ノードを構成する文書番号毎に、その文書番号に対応する図5のデータ構造を有するメタインデックス304のエントリを抽出し、そのエントリから、作者名、行数、色番号、及びタイトルを抽出する（ステップ1403）。これらの抽出されたデータは、上記各ノードに対応付けて記憶される。

【0100】

次に、表示装置306は、ステップ1401で読み込んだ文書数と、ステップ1403で抽出した各ノードの色番号とから、スレッドビューの先頭で表示される各タイトルの色を決定する（ステップ1404）。この動作は、各色番号に実際の色をマッピングする動作として実現される。

【0101】

次に、表示装置306は、スレッドに含まれるルート文書のタイトルとその他

のタイトルを、"Main Topic:" 及び"Sub Topic:" に続けて表示するためのHTML文書を作成する。この場合に、各タイトルは、前述した構造解析部402が管理する図12に示されるカラーテーブルの上記スレッドに属する各エントリから順次読み出され、同時に順次読み出される各色番号からステップ1404で決定された各色が算出され、その各色での表示が順次指定される。各色は、HTML文書の色指定命令 (タグ等) によって指定される。

【0102】

最後に、表示装置306は、ステップ1402で復元したスレッドのツリー構造を示す配列データを構成する左端のノードの文書番号から順に処理することにより、そのツリー構造を表示するためのHTML文書を作成する(ステップ1406)。この場合、前述したように、表示装置306は、ステップ1403で抽出した各ノードの作者名、行数、及び色番号に基づいて、ツリー構造の各ノードの文書番号を、そのノードに対応する文書の作者名の先頭文字(2バイト)とその文書の行数とからなる表示データ、

作者名 [行数]

に変換し、更に、その表示データをそのノードの色番号に対応する色で表示させるためのHTML文書データを生成する。色番号と実際の色との対応関係は、ステップ1404で決定された対応関係に従う。また、前述したように、表示装置306は、各ノードに対応する上記表示データの前後に、その接続関係に基づいて、" * "、" + "、" = "、又は" . " 等の記号を表示するためのHTML文書データを生成する。ここで、ツリー構造をそのままの形式で表示可能とするために、例えば、HTMLにおける制御用タグであるプリフォーマットタグ <PRE> が使用される。更に、上記ノード毎の表示データは、そのノードに対応する文書群データベース301内の文書データを表示するためのアプリケーションへのURLがHREF属性の値として指定されるアンカータグとして生成される。

【0103】

続いて、表示装置306は、スレッドインデックス305のファイルから文書末尾(EOF)を検出するまで、上記ステップ1401~1406の一連の処理

を繰り返し実行することにより、各スレッド毎のビューデータを作成する（ステップ1407→1401）。

【0104】

表示装置306は、スレッドインデックス305のファイルから文書末尾を検出すると（ステップ1407の判定がYES）、最終的に得られたHTMLテーブル文書データをWebサーバに引き渡して、スレッドビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図26に例示されるような形式で、スレッドビューが表示される。

【0105】

次に、表示装置306が実現する発言者ビューの表示動作につき説明する。

前述したように、発言者ビューにおいては、各文書のタイトルが、発言者（作者）毎に分類され、かつ発言者が発言の多い順にソートされ、同一発言者内では日付順で、図27に示される表示形態で、表示される。

【0106】

図16は、表示装置306が実行する発言者ビューの表示動作を示す動作フローチャートである。。

表示装置306は、発言者ビューを実現するために、図17のデータ構造を有する作者配列データを使用する。そして、表示装置306は、発言者ビューの表示開始時に、この作者配列データを初期化する（ステップ1601）。

【0107】

次に、表示装置306は、図5のデータ構造を有するメタインデックス304のファイルから1つのエントリのデータを読み込む（ステップ1602）。

次に、表示装置306は、このエントリから抽出される作者名の作者が、作者配列データに含まれていない作者であるか否かを判定する（ステップ1603）。

【0108】

表示装置306は、上記エントリから抽出される作者名の作者が、作者配列データに含まれていない作者である場合には、作者配列データに新しい作者項目を追加する（ステップ1603→1604）。表示装置306は、上記エントリか

ら抽出される作者名の作者が、作者配列データに含まれている作者である場合には、ステップ1604の処理は実行しない。

【0109】

次に、表示装置306は、作者配列データ中の該当する作者項目に、上記エントリから抽出される文書番号を登録する（ステップ1605）。

続いて、表示装置306は、メタインデックス304のファイルから文書末尾（EOF）を検出するまで、上記ステップ1602～1605の一連の処理を繰り返し実行することにより、メタインデックス304に登録されている全ての文書番号を、作者別に作者配列データに登録する。

【0110】

表示装置306は、メタインデックス304のファイルから文書末尾を検出すると（ステップ1606の判定がNO）、作者配列データ中の各作者項目を、それぞれの項目に登録されている文書番号の数、即ち各作者毎の発言文書数に基づいてソートする（ステップ1607）。

【0111】

続いて、表示装置306は、作者配列データ中の同一作者項目内で、文書番号を、それに対応するメタインデックス304中のエントリから抽出される日付に基づいてソートする（ステップ1608）。

【0112】

最後に、表示装置306は、上記ステップ1607及び1608でのソートの結果得られる作者配列データの各作者項目毎に、作者名と、その項目内の各文書番号に対応するメタインデックス304中のエントリから抽出される日付及びタイトルを表示するためのHTMLテーブル文書データを生成し、それをWebサーバに引き渡して、発言者ビューの表示動作を終了する。この結果、ユーザの端末のWebブラウザアプリケーション上に、図27に例示されるようなテーブル形式で、発言者ビューが表示される。

【0113】

ユーザは、発言者ビュー上で、発言者及び発言日付という観点から、文書集合（会議室）内の文書を参照することが可能となる。

また、或る発言者の発言を時間を追って参照したり、会議室内で多くの発言をするリーダー的な発言者を一目で確認することができる。

表示装置306の他の表示態様

次に、上記各ビューの表示動作以外に表示装置306が実現する各表示動作の態様について説明する。

【0114】

まず、表示装置306が実現する発言内容表示の動作につき説明する。

前述したように、ユーザは、キーワードビューにおけるそれぞれのスレッド上のアンカー又はスレッドビューにおける各ノード上のアンカーを、マウス装置等でクリックすることにより、各スレッドのルート文書又は各ノードに対応する文書等に、即座にアクセスすることができる。

【0115】

ユーザによってこれらの操作が実行された場合には、Webサーバから指示によって、表示装置306によって実行される図18に示される動作フローチャートの処理が例えばCGIとして起動される。この場合、この処理には、ユーザによって指定されたアンカータグに含まれる文書番号の情報が引き渡される。

【0116】

この結果まず、表示装置306は、上記文書番号の情報を読み込んだ後（ステップ1801）、ヘッダ部に上記読み込んだ文書番号と同じ文書番号を含んでいる文書ファイルを読み込むまで、文書群データベース301からの文書ファイルの読み込みを行う（ステップ1802→1803→1802のループ）。

【0117】

表示装置306は、ヘッダ部に上記読み込んだ文書番号と同じ文書番号を含んでいる文書ファイルを読み込むと（ステップ1803の判定がYES）、新しい文書のヘッダ部を読み込むまで、ステップ1804～1809のループにより、上記文書ファイルから1行ずつデータを読み込み、そのデータを1行分のHTML文書データに変換し、そのHTML文書データをWebサーバに出力する（ステップ1808）。

【0118】

この場合に、各行のデータが他の文書等へのURLを含んでいる場合には、表示装置306は、そのデータを上記URLがHREF属性の値として指定されるアンカータグに変換した上で出力する（ステップ1804→1805）。

【0119】

この結果、ユーザは、発言内容の表示中のアンカーを更にマウス装置等によってクリックすることにより、更に他のリソースにジャンプすることができる。

また、各行のデータが他の文書の行を引用したコメント行である場合には、表示装置306は、そのデータの色を変換するタグを追加した上で出力する（ステップ1806→1807）。

【0120】

この結果、ユーザは、コメント行を一目で判別することができる。

表示装置306は、該当する文書データの出力処理を終了すると、上記文書を含むスレッドのツリー構造を表示するHTML文書データを生成し出力して、発言内容表示の動作を終了する（ステップ1809→1810）。この処理は、前述した図14の動作フローチャートで示されるスレッドビューの表示動作と同様にして実現できる。

【0121】

以上の表示動作の結果、ユーザの端末のWebブラウザアプリケーション上には、例えば図28に示されるように、表示画面の上半分に発言内容が表示され、表示画面の下半分にはその発言内容の文書を含むスレッドのツリー構造が表示される。なお、この表示画面には、図28に示されるように、キーワードビューやスレッドビューを表示させるためのアンカーや、検索を実行するためのアンカーと等を同時に表示させることもできる。

【0122】

これらのビューの切替え機能により、例えば、キーワードビュー → スレッドビュー → 発言内容表示 → 発言者ビュー 発言内容表示 → スレッドビュー → … というように、会議室内の文書（発言）をユーザの嗜好に応じて横断的に参照してゆくことが可能となる。

【0123】

次に、表示装置306が実現する作者別／日付別色分け表示の動作につき説明する。図19は、その動作を示す動作フローチャートである。

まず、表示装置306は、メタインデックス304及びスレッドインデックス305に基づいて、図20(a)に示されるように作者項目毎に文書番号が分類された作者配列データと、図20(b)に示されるように日付項目毎に文書番号が分類された日付配列データとを予め作成する。これらの作成処理の詳細は省略するが、前述した図16の動作フローチャートと同様の処理によって実現できる。そして、作者配列データ中の各作者項目又は日付配列データ中の各日付項目に、それぞれ異なる色が割り当てられる。この色の割当ては、作者項目毎の作者の総数又は日付項目毎の日付の総数から決定される。

【0124】

次に、表示装置306は、ユーザの指定に基づく項目選択ボタン情報をWebサーバを経由して取得し、作者ボタンが押されたか日付選択ボタンが押されたかを判定する（ステップ1902、1904、図29参照）。

【0125】

表示装置306は、作者ボタンが押されたと判定した場合には、図20(a)に示される作者配列データを参照することにより、スレッドツリーの表示データを作成し出力する（ステップ1902→1903）。この処理は、前述した図14の動作フローチャートと同様の処理によって実現されるが、この場合に、ツリーの各ノードは、そのノードに対応する作者名に対応する作者配列データ中の作者項目に割り当てられている色で表示される。

【0126】

一方、表示装置306は、日付選択ボタンが押されたと判定した場合には、図20(b)に示される日付配列データを参照することにより、スレッドツリーの表示データを作成し出力する（ステップ1904→1905）。この処理も、前述した図14の動作フローチャートと同様の処理によって実現されるが、この場合に、ツリーの各ノードは、そのノードに対応する日付に対応する日付配列データ中の日付項目に割り当てられている色で表示される。

【0127】

以上の表示動作の結果、例えばユーザが作者ボタンを押した場合には、ユーザの端末のWebブラウザアプリケーション上には、例えば図29に示されるように、作者別に色分けされたスレッドのツリーが表示され、ユーザは同一の作者の文書を一目で確認することができる。

【 0 1 2 8 】

次に、表示装置306が実現するスレッドビューを使った検索結果の強調表示の動作につき説明する。図21はその動作を示す動作フローチャートである。

まず、表示装置306は、検索後入力フォーム画面を表示するためのHTML文書データを生成し出力する（ステップ2101）。この結果、ユーザの端末のWebブラウザアプリケーション上には、例えば図30に示されるような検索入力フォーム画面が表示される。ユーザは、この検索入力フォームに検索語を入力して検索の実行を指定する。

【 0 1 2 9 】

上記検索入力フォームに入力された検索語は、Webサーバを経由して文字列検索装置405（図4）に引き渡される。文字列検索装置405は、ユーザによる検索語の指定に基づいて、索引ファイル404を用いながら文書群データベース301内の指定されたスレッドを構成する各文書に対して全文検索を実行し、その検索語を含む文書番号を出力する（ステップ2102、2103）。

【 0 1 3 0 】

表示装置306は、上記検索語を含む文書番号を受け取ると、その文書番号を含むスレッドのツリー構造を表示するHTML文書データを、前述した図14の動作フローチャートと同様の処理によって表示する。この場合に、表示装置306は、上記文書番号を含むノードの色を強調色に指定する（ステップ2104、2105）。

【 0 1 3 1 】

この結果、ユーザの端末のWebブラウザアプリケーション上において、例えば図31に示されるような検索結果に基づくスレッドビューの強調表示が実現される。これにより、ユーザは、スレッドの構造を把握しつつ、検索を実行することができる。

【 0 1 3 2 】

次に表示装置 3 0 6 が実現するキーワードビューを使った検索結果の強調表示の動作につき説明する。図 2 2 はその動作を示す動作フローチャートである。

まず、表示装置 3 0 6 は、図 2 1 のステップ 2 1 0 1 の場合と同様に、検索後入力フォーム画面を表示するための HTML 文書データを生成し出力する（ステップ 2 2 0 1）。ユーザは、この検索入力フォームに検索語を入力して検索の実行を指定する。

【 0 1 3 3 】

上記検索入力フォームに入力された検索語は、Webサーバを經由して文字列検索装置 4 0 5（図 4）に引き渡される。文字列検索装置 4 0 5 は、ユーザによる検索語の指定に基づいて、索引ファイル 4 0 4 を用いながら文書群データベース 3 0 1 内の指定された会議室を構成する各文書に対して全文検索を実行し、その検索語を含む文書番号を出力する（ステップ 2 2 0 2、2 2 0 3）。

【 0 1 3 4 】

表示装置 3 0 6 は、上記検索語を含む文書番号を受け取ると、まず、図 6 のデータ構造を有するスレッドインデックス 3 0 5 を参照して、上記文書番号を含むエントリに対応するルート文書番号を抽出する（ステップ 2 2 0 4）。

【 0 1 3 5 】

続いて、表示装置 3 0 6 は、指定された会議室に関するキーワードビューを表示する HTML 文書データを、前述した図 1 3 の動作フローチャートと同様の処理によって表示する。この場合に、表示装置 3 0 6 は、ステップ 2 2 0 4 で抽出されたルート文書番号に対応するスレッドのタイトル又はその表示エリア全体の色を強調色に指定し、更に、表示されるキーワード中に検索語が含まれている場合には、そのキーワードも強調色に指定する（ステップ 2 2 0 5、2 2 0 6、2 2 0 7）。

【 0 1 3 6 】

この結果、ユーザの端末の Web ブラウザアプリケーション上において、例えば図 3 2 に示されるような検索結果に基づくキーワードビューの強調表示が実現される。これにより、ユーザは、検索語を含むスレッドを一目で把握することが

できる。

【 0 1 3 7 】

なお、表示装置 3 0 6 は、検索結果の文書番号とそれに対応するタイトルを、例えば図 3 3 に示されるように羅列して表示するように構成することも可能である。

【 0 1 3 8 】

最後に、表示装置 3 0 6 が実現するサブピック毎のキーワードビューの表示動作について説明する。

前述したキーワードビューは、スレッド毎にキーワードを表示するものであった。これに対して、サブピック毎のキーワードビューでは、1つのスレッド内のサブピック毎に、キーワードを抽出して表示することができる。

【 0 1 3 9 】

この動作において、表示装置 3 0 6 は、図 2 4 のデータ構造を有するサブピックインデックスを使用する。サブピックインデックスは、図 1 2 に示されるカラーテーブルのデータ構造に対して、キーワードフィールドが追加されたデータ構造を有する。

【 0 1 4 0 】

サブピックインデックスは、実質的には前述したカラーテーブルを置き換えるものであるため、サブピックインデックスにおけるキーワードフィールド以外のフィールドの内容は、構造解析部 4 0 2 による前述した図 8 のステップ 8 1 0 の処理によって予め登録されている。この場合、図 8 のステップ 8 1 0 の処理の説明において前述したように、カラーテーブルであるサブピックインデックスには、ルート文書番号毎（スレッド毎）に、それに含まれるルート文書のタイトルを示すメインピックと、それ以外の文書のタイトルを示すサブピックのそれぞれに対応するエントリが得られる。表示装置 3 0 6 は、この登録内容を利用する。

【 0 1 4 1 】

図 2 3 は、表示装置 3 0 6 が実現するサブピックからのキーワード抽出の制御を示す動作フローチャートである。

まず、表示装置 3 0 6 は、各スレッドについて、サブトピックインデックス内のそのスレッドに含まれる各エントリに登録されている文書番号に基づいて、メイントピック及びサブトピック単位で、それぞれに属する文書群を各結合文書ファイルにまとめ（ステップ 2 3 0 1）、その結果得られる各結合文書ファイルの内容解析部 4 0 3（図 4）に入力する（ステップ 2 3 0 2）。

【 0 1 4 2 】

内容解析部 4 0 3 は、各結合文書ファイル別にキーワードを抽出し、その結果を表示装置 3 0 6 に返す。表示装置 3 0 6 は、内容解析部 4 0 3 から返された各結合文書ファイル別のキーワードを、サブトピックインデックス内の上記各結合文書ファイルに対応するエントリのキーワードフィールドに登録する（ステップ 2 3 0 3）。

【 0 1 4 3 】

以上のようにして、各スレッドについて、メイントピック及びサブトピック単位で、それぞれに属する文書群からキーワードが抽出される。

その後は、表示装置 3 0 6 は、サブトピックインデックスの内容に基づいて、ユーザにより指定されたスレッドに関して、そのスレッドのメイントピック及びサブトピック単位で、それぞれのタイトルとそれぞれに属するキーワードを表示するための HTML 文書データを生成し出力する。

【 0 1 4 4 】

この結果、ユーザの端末の Web ブラウザアプリケーション上には、例えば、図 3 4 に示されるような形式で、サブトピック毎のキーワードビューが表示される。これにより、ユーザは、キーワードによるより精密なトピックの絞込みを行うことができる。

本実施の形態を実現するプログラムが記録された記録媒体についての補足

本発明は、計算機により使用されたときに、上述の本発明の実施の形態の各構成によって実現される機能と同様の機能を計算機に行わせるための計算機読出し可能記憶媒体として構成することもできる。

【 0 1 4 5 】

この場合に、図 3 5 に示されるように、例えばフロッピーディスク、CD-R

OMディスク、光ディスク、リムーバブルハードディスク等の可搬型記憶媒体 3 5 0 2 や、ネットワーク回線 3 5 0 3 経由で、本発明の実施の形態の各種機能を実現するプログラムが、コンピュータ 3 5 0 1 の本体 3 5 0 4 内のメモリ (R A M又はハードディスク等) 3 5 0 5 にロードされて、実行される。

【 0 1 4 6 】

【発明の効果】

本発明の第 1 の態様の構成によれば、キーワードを頼りにして、雑多な文書集合の中から必要な文書が含まれているであろう文書群を正確かつ容易に見つけ出すことが可能となる。

【 0 1 4 7 】

本発明の第 2 の態様の構成によれば、検索指定に該当する文書群を一目で把握することが可能となる。

本発明の第 3 の態様の構成によれば、文書群全体の構造を容易に把握することが可能となる。

【 0 1 4 8 】

本発明の第 4 の態様の構成によれば、文書群全体の話題の推移を容易に把握することが可能となる。

本発明の第 5 の態様の構成によれば、スレッドの構造を把握しつつ、検索を実行することが可能となる。

【 0 1 4 9 】

本発明の第 6 の態様の構成によれば、キーワードによるより精密なトピックの絞込みを行うことが可能となる。

このように、本発明によれば、文書群に対して種々の観点からアクセスすることが可能となる。

【図面の簡単な説明】

【図 1】

参照関係を有する文書群の例を示す図である。

【図 2】

文書のデータ構造の例を示す図である。

【図 3】

本発明の実施の形態のシステム構成図（その 1）である。

【図 4】

本発明の実施の形態のシステム構成図（その 2）である。

【図 5】

メタインデックスのデータ構造を示す図である。

【図 6】

スレッドインデックスのデータ構造を示す図である。

【図 7】

索引ファイルのデータ構造を示す図である。

【図 8】

書式解析部と構造解析部の動作フローチャートである。

【図 9】

文書番号のスレッドインデックスへの登録の動作フローチャートである。

【図 10】

文書番号のスレッドインデックスへの登録の動作説明図である。

【図 11】

色番号登録の動作フローチャートである。

【図 12】

カラーテーブルの例を示す図である。

【図 13】

キーワードビューの動作フローチャートである。

【図 14】

スレッドビューの動作フローチャートである。

【図 15】

スレッドビューの制御用配列の例を示す図である。

【図 16】

発言者ビューの動作フローチャートである。

【図 17】

発言者ビューの制御用配列の例を示す図である。

【図18】

発言内容表示の動作フローチャートである。

【図19】

作者別／日付別色分け表示の動作フローチャートである。

【図20】

作者別／日付別色分け表示用配列の例を示す図である。

【図21】

スレッドビューを使った検索結果の強調表示の動作フローチャートである。

【図22】

キーワードビューを使った検索結果の強調表示の動作フローチャートである。

【図23】

サブピックからのキーワード抽出の制御動作フローチャートである。

【図24】

サブピックインデックスの例を示す図である。

【図25】

キーワードビューの表示例を示す図である。

【図26】

スレッドビューの表示例を示す図である。

【図27】

発言者ビューの表示例を示す図である。

【図28】

発言内容の表示例を示す図である。

【図29】

スレッドビューを用いた文書属性「作者」の強調（色別）表示の例を示す図である。

【図30】

会議室内の全発言の検索（入力）表示例を示す図である。

【図31】

スレッドビューを用いた文字列「プロトコル」を含むノードの強調表示の例を示す図である。

【図 3 2】

キーワードビューを用いた文字列「プロトコル」を含むスレッドの強調表示の例を示す図である。

【図 3 3】

会議室内の全発言の検索（結果出力）表示例を示す図である。

【図 3 4】

サブピックから抽出したキーワードの表示例を示す図である。

【図 3 5】

本実施の形態を実現するプログラムが記録された記録媒体の説明図である。

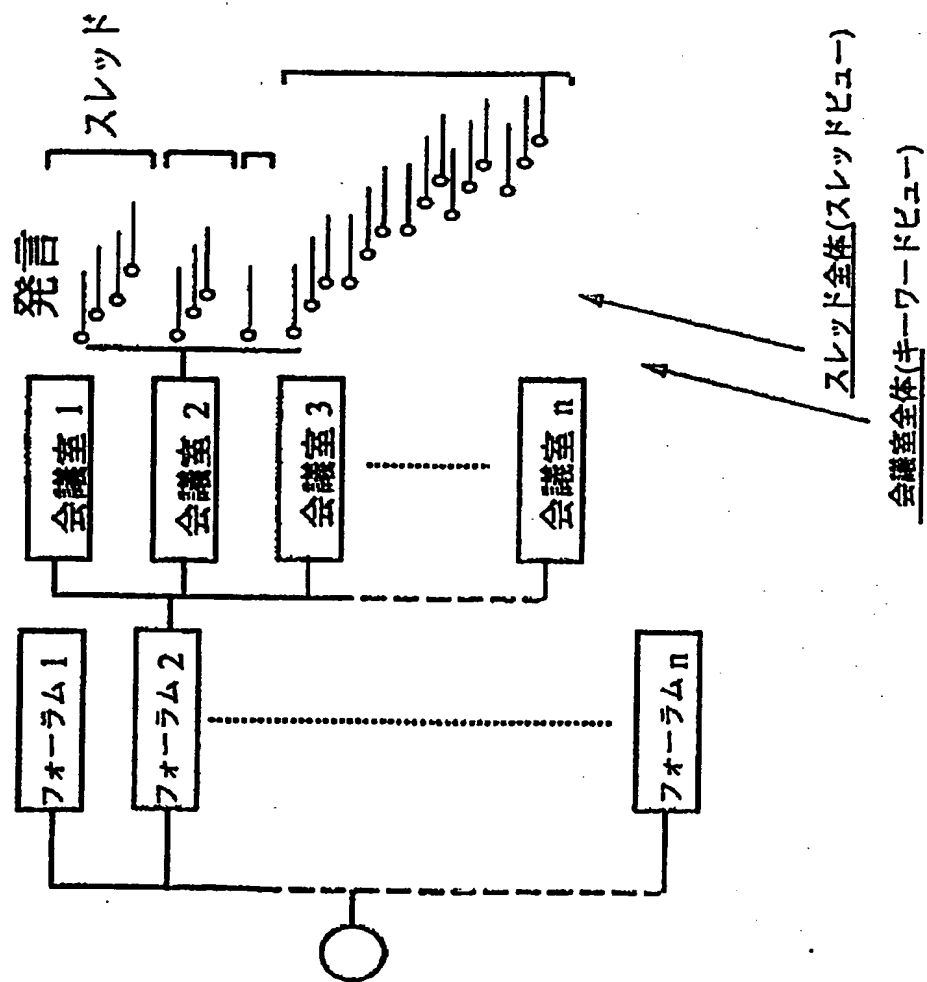
【符号の説明】

- | | |
|-------|------------|
| 3 0 1 | 文書群データベース |
| 3 0 2 | 文書群解析装置 |
| 3 0 3 | 集計装置 |
| 3 0 4 | メタインデックス |
| 3 0 5 | スレッドインデックス |
| 3 0 6 | 表示装置 |
| 4 0 1 | 書式解析部 |
| 4 0 2 | 構造解析部 |
| 4 0 3 | 内容解析部 |
| 4 0 4 | 索引ファイル |
| 4 0 5 | 文字列検索装置 |

【書類名】 図面

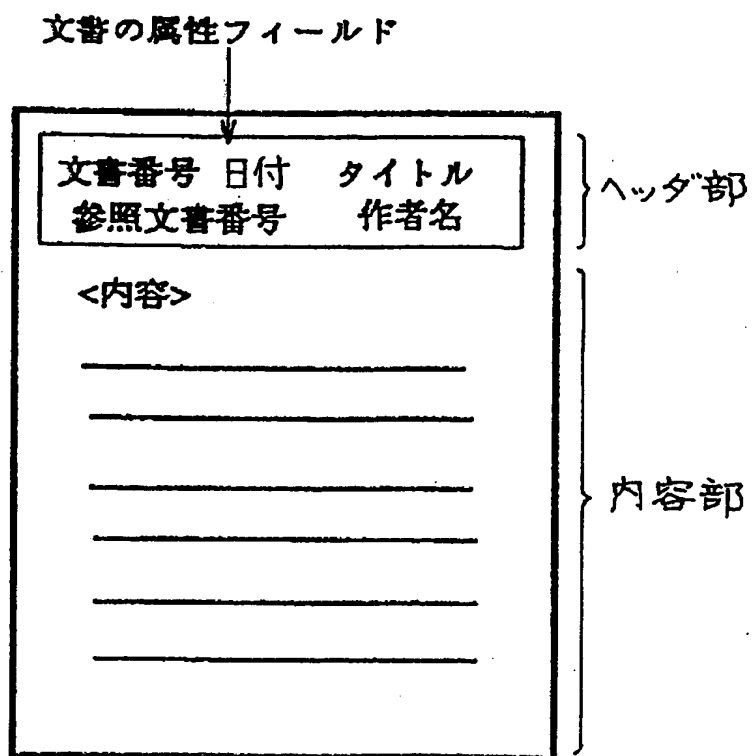
【図1】

参照関係を有する文書群の例を示す図



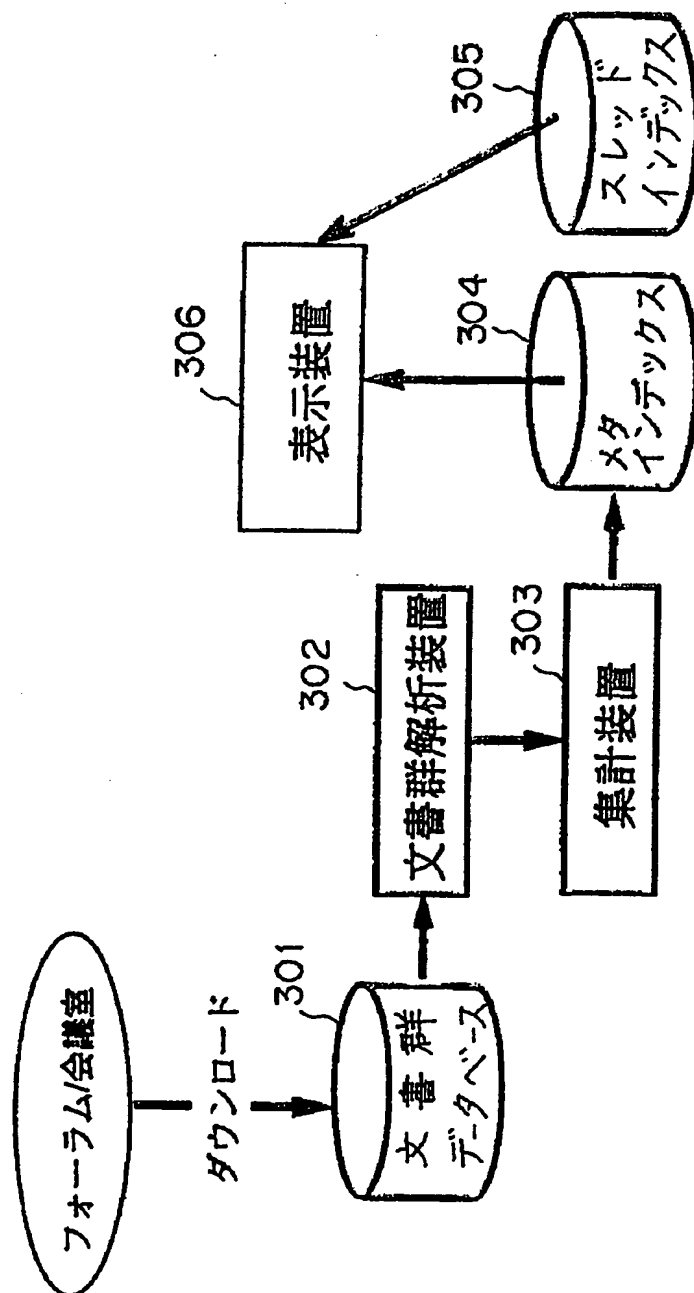
【図2】

文書のデータ構造の例を示す図



【図3】

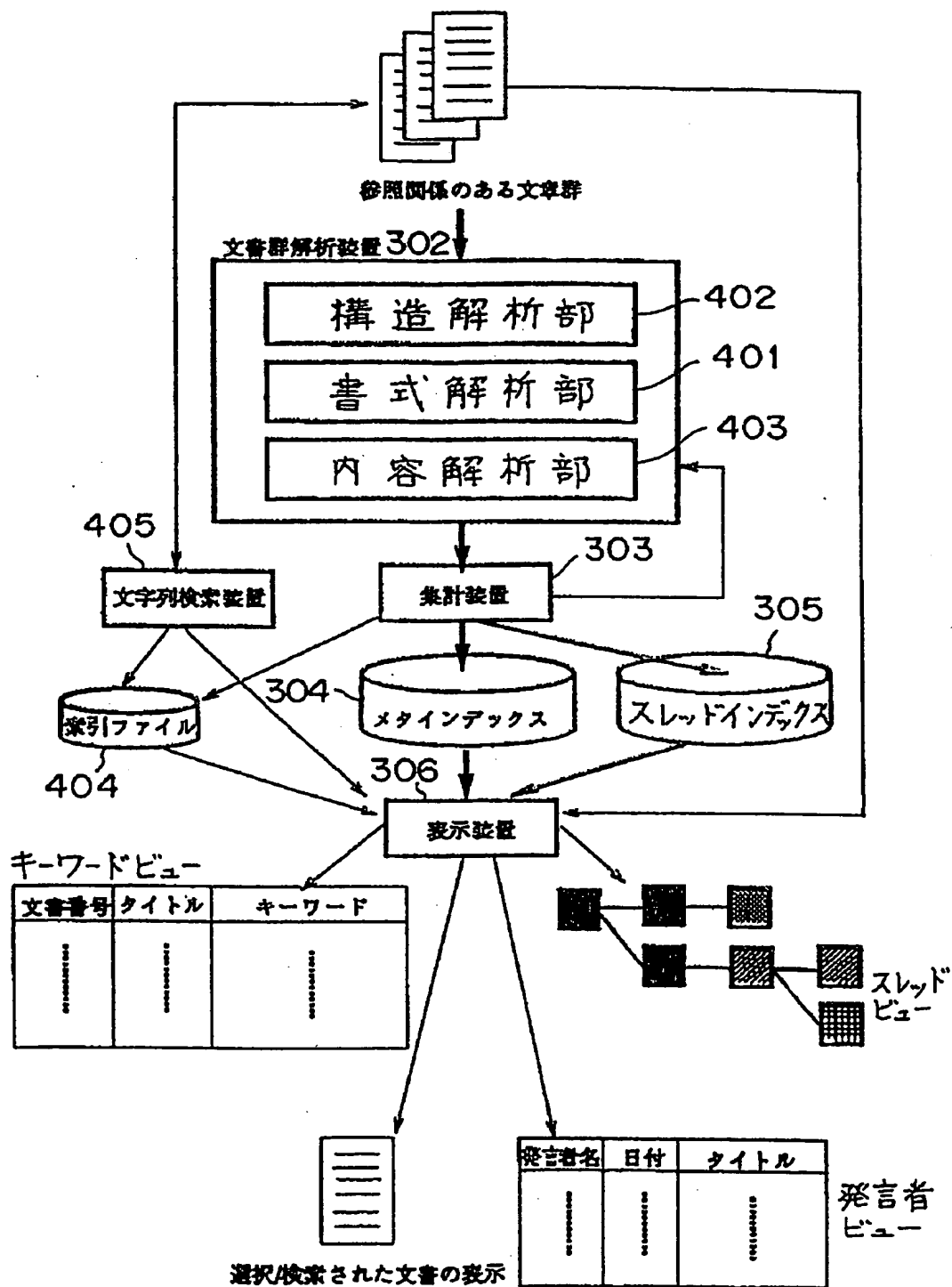
本発明の実施の形態のシステム構成図
(その1)



【図4】

本発明の実施の形態のシステム構成図

(その2)



【図5】

メタインデックスのデータ構造を示す図

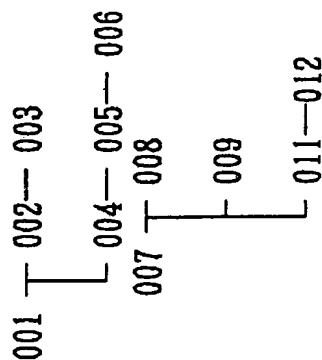
文書番号	参照文書番号	タイトル	作者名	キーワード	行数	色番号	日付
004	001	改えて	内野	—	12	4	96/10/05
005	004	Re:改えて	津田	—	43	4	96/10/08
006	005	Javaのアプリ	内野	—	80	6	96/10/15
*007	—	HD不良?	松井	SunOS, HD, 64MB, フォーマット	12	1	96/10/18
—	—	—	—	—	—	—	—

*はルートの文書番号

【図 6】

スレッドインデックスのデータ構造を示す図

ルート 文書番号	文書数	スレッドの構造
∴		
001	6	(001(002 003)(004(005 006)))
007	5	(007 008 009(011 012))
∴ ∴ ∴		



リストの意味：（親文書番号 子文書番号／サブツリー 子文書番号／サブツリー.....）

【図7】

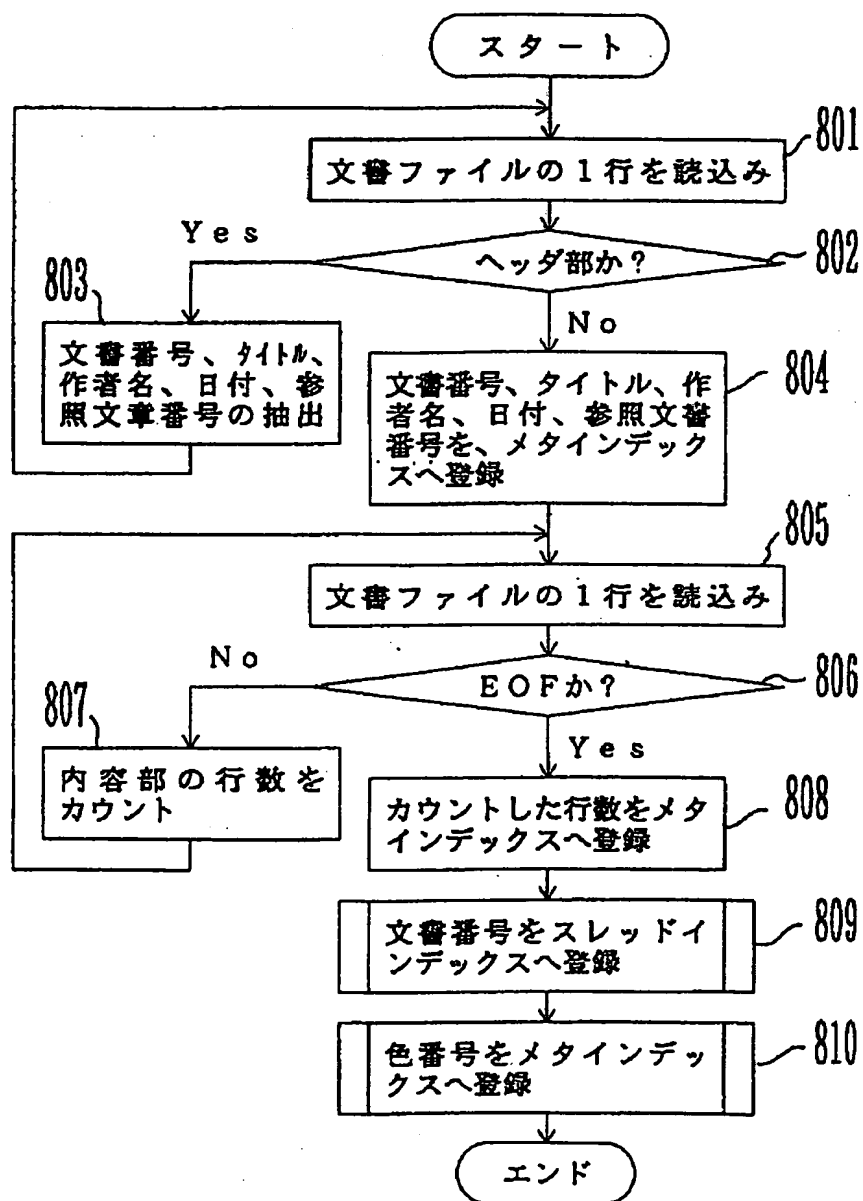
索引ファイルのデータ構造を示す図

索引語	索引語を含むキーワード	スレッド番号*
インターネット	・インターネットスタート ・インターネットへの接続	00496 00218
コントロール	・コントロールパネル	00518
ドメイン	・ドメインネームサーバ ・国のドメイン	00128 00476
⋮	⋮	⋮

*スレッド番号はスレッドのルートの文書番号

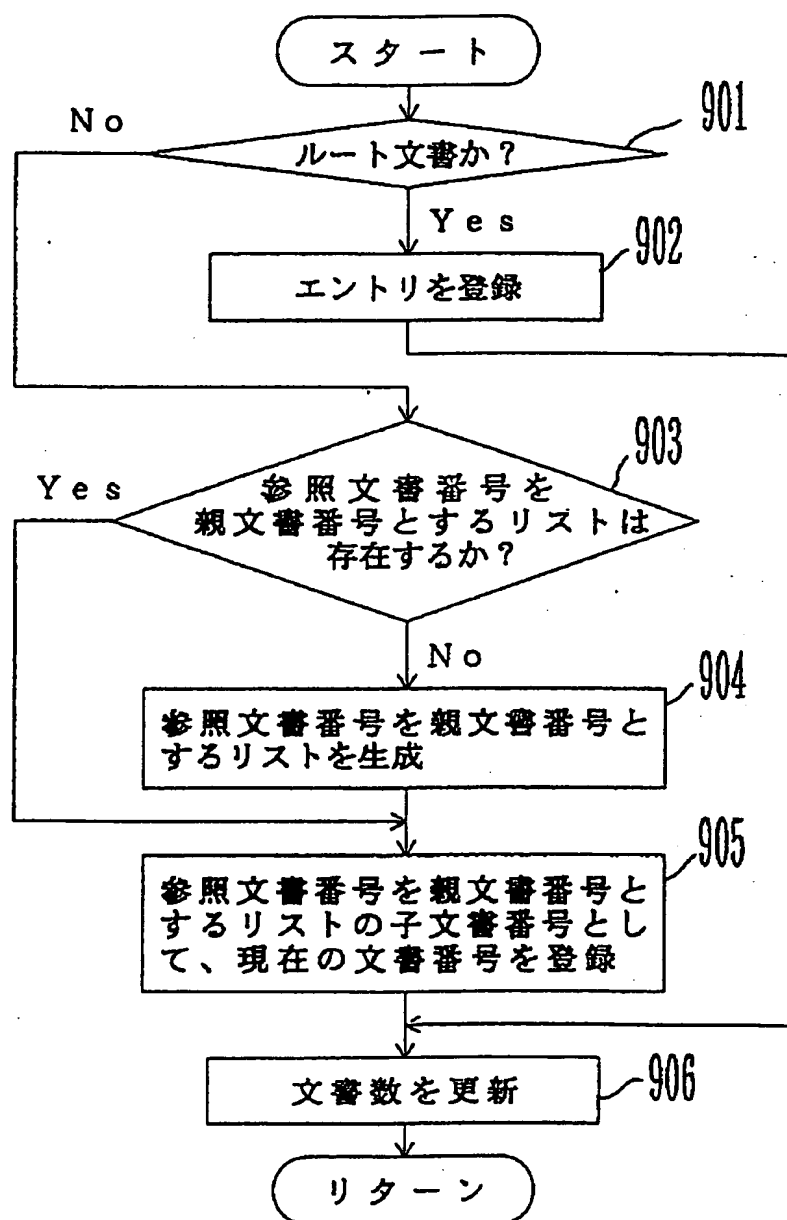
【図8】

書式解析部と構造解析部の動作フローチャート



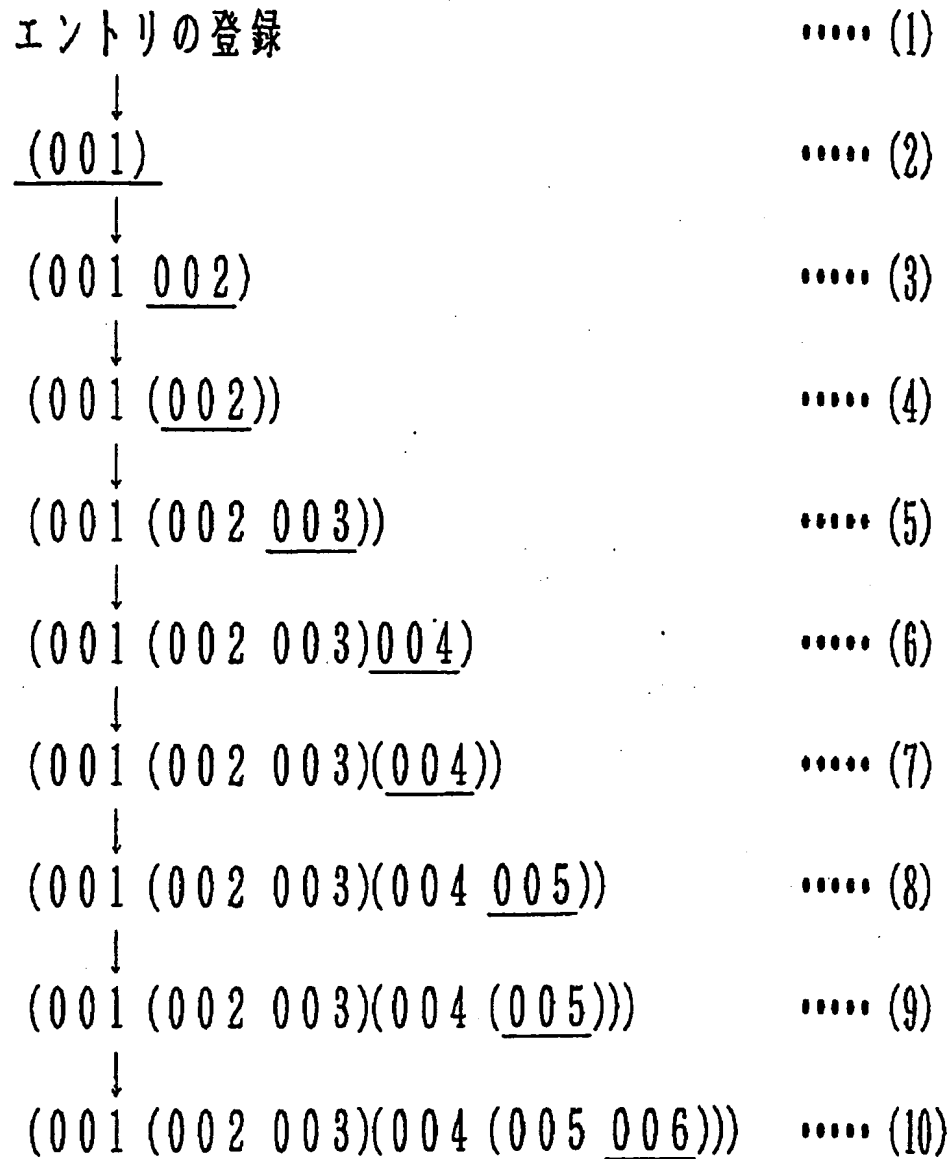
【図9】

文書番号のスレッドインデックスへの登録の動作フローチャート



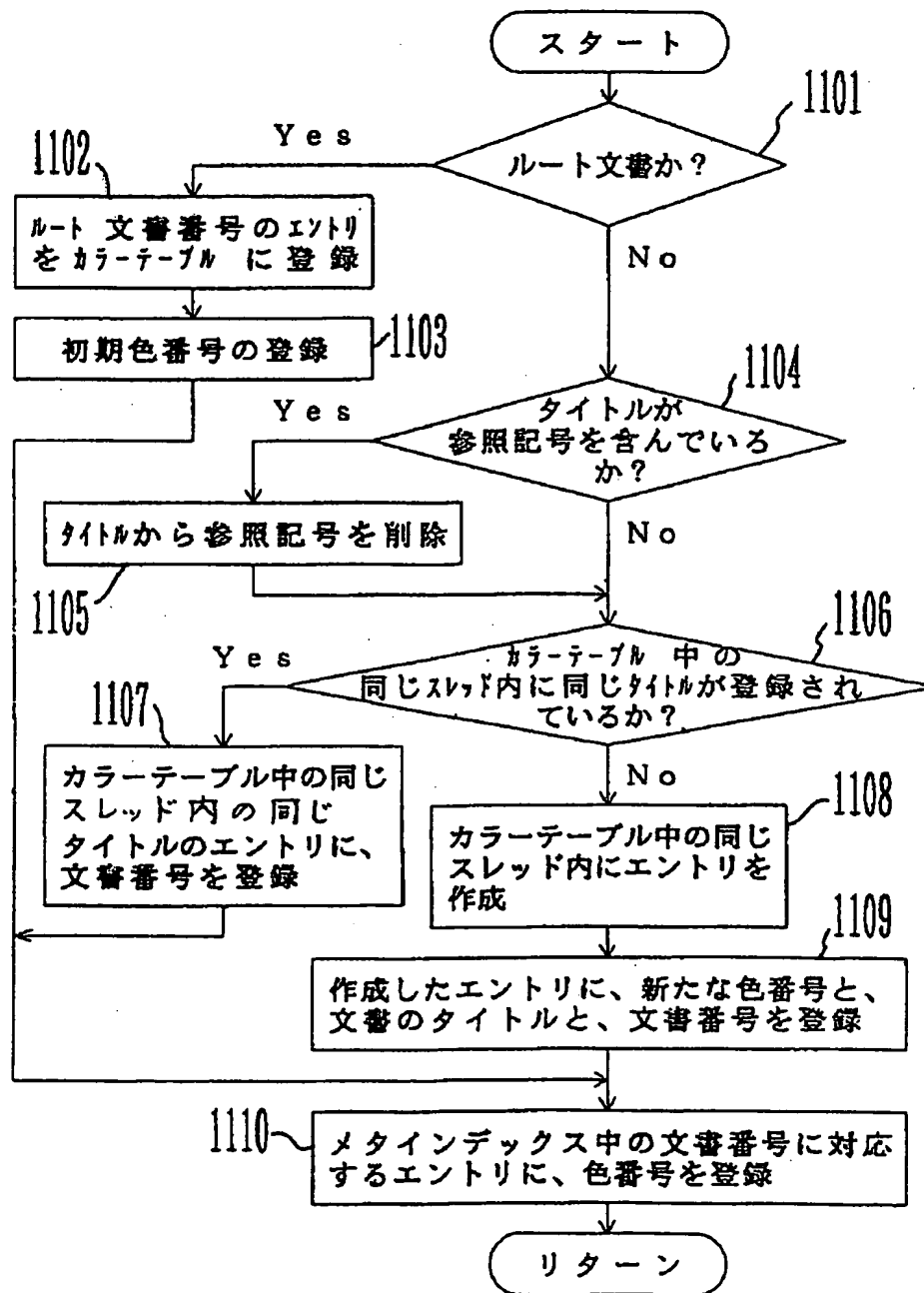
【図 1 0】

文書番号のスレッドインデックスへの登録の動作説明図



【図11】

色番号登録の動作フローチャート



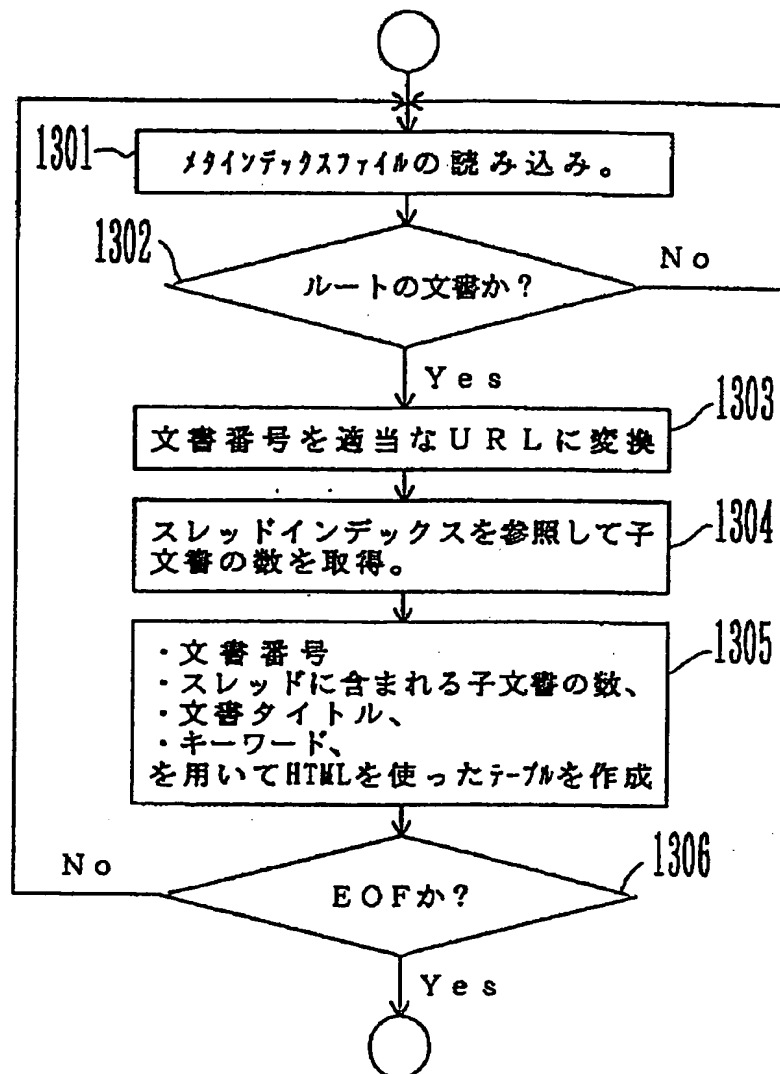
【図 1 2】

カラーテーブルの例を示す図

ルート 文書番号	色番号	タイトル	文書番号
⋮	⋮	⋮	⋮
001	#1	(メインピク)	001, 002, 004
	#2	(サブピク 1)	003
	#3	(サブピク 2)	005, 006
⋮	⋮	⋮	⋮

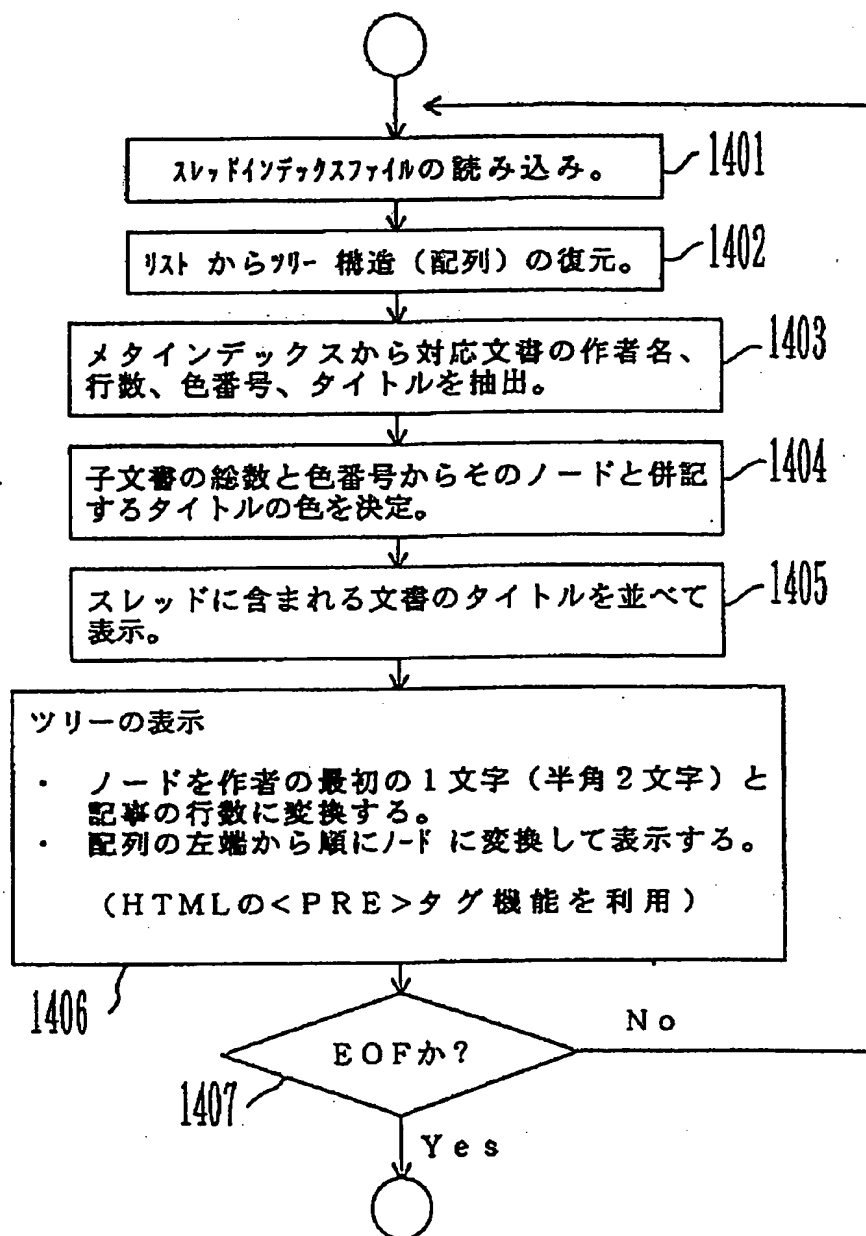
【図13】

キーワードビューの動作フローチャート



【図 1 4】

スレッドビューの動作フローチャート



【図 1 5】

スレッドビューの制御用配列の例を示す図

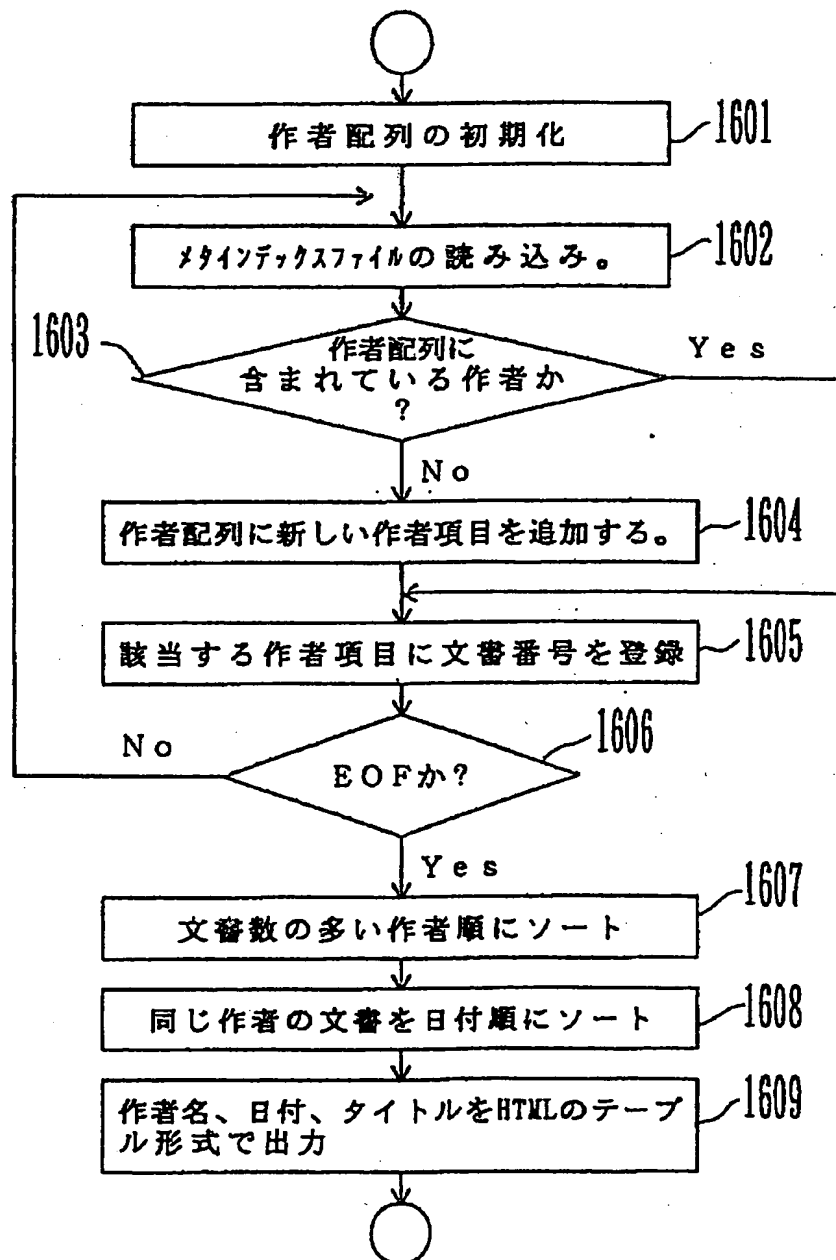
ツリー構造のリスト : (001(002 003)(004(005 006)))

	1	2	3	4
1	001	002	003	空 白
2	空 白	004	005	006

ツリー構造を表す配列

【図 1 6】

発信者ビューの動作フローチャート



【図 1 7】

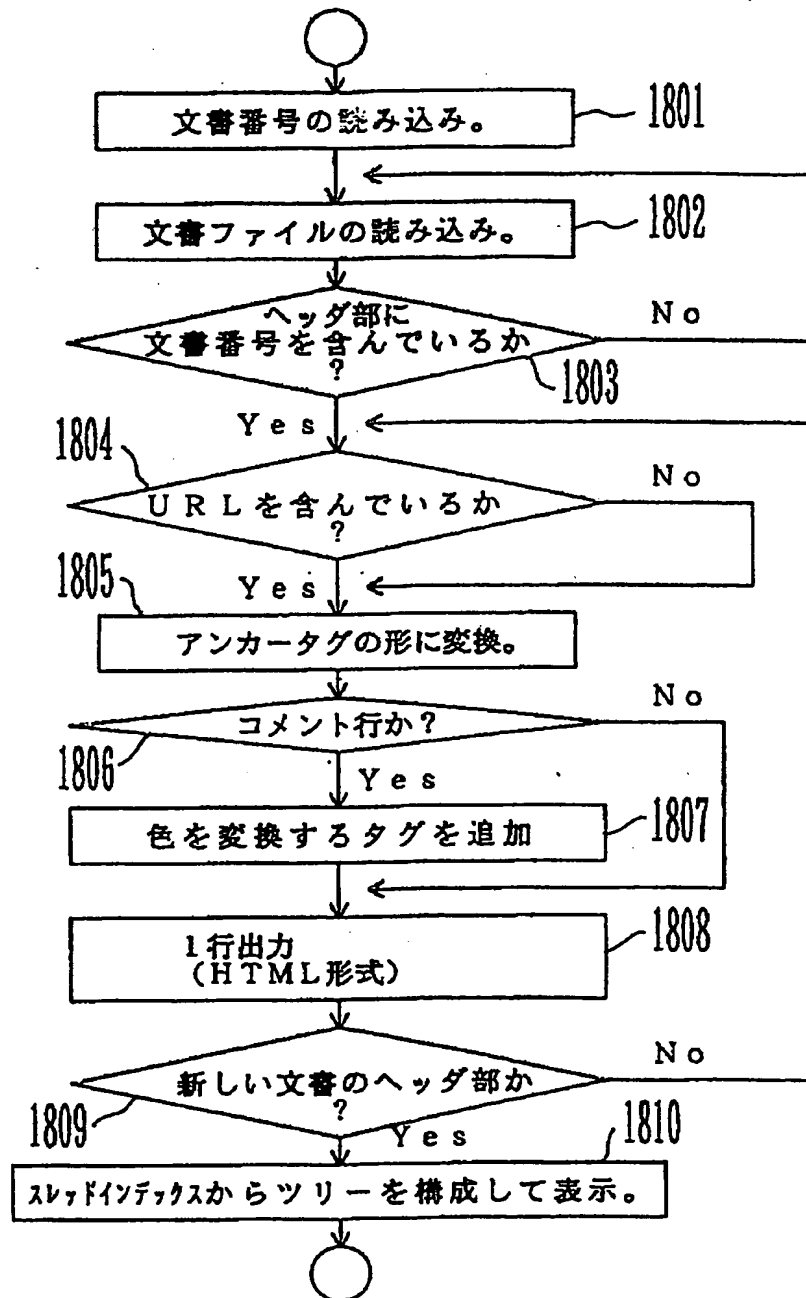
発言者ビューの制御用配列の例を示す図

作者配列の例

作者 A	001, 005, 007, 008
作者 B	011, 012, 013
作者 C	003, 009, 010
作者 D	004, 006
作者 E	002

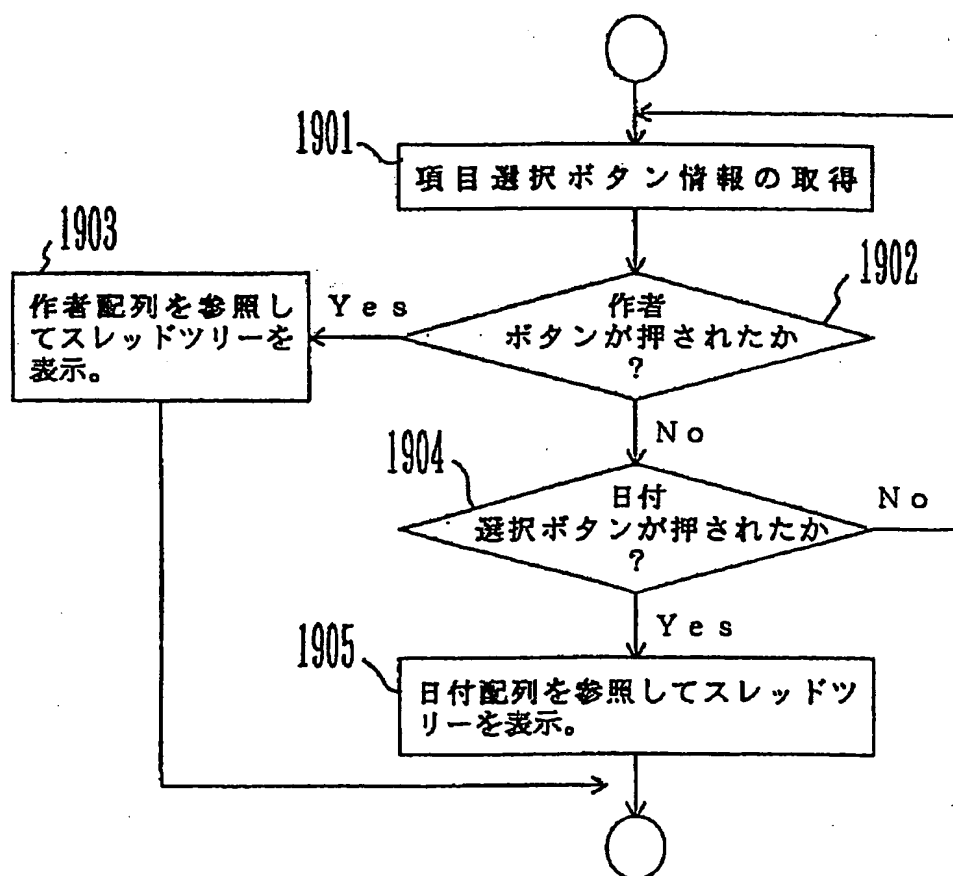
【図18】

発言内容表示の動作フローチャート



【図 1 9】

作者別 / 日付別色分け表示の動作フローチャート



【図 2 0】

作者別 日付別

色分け表示用配列の例を示す図

(b) 日付配列の例

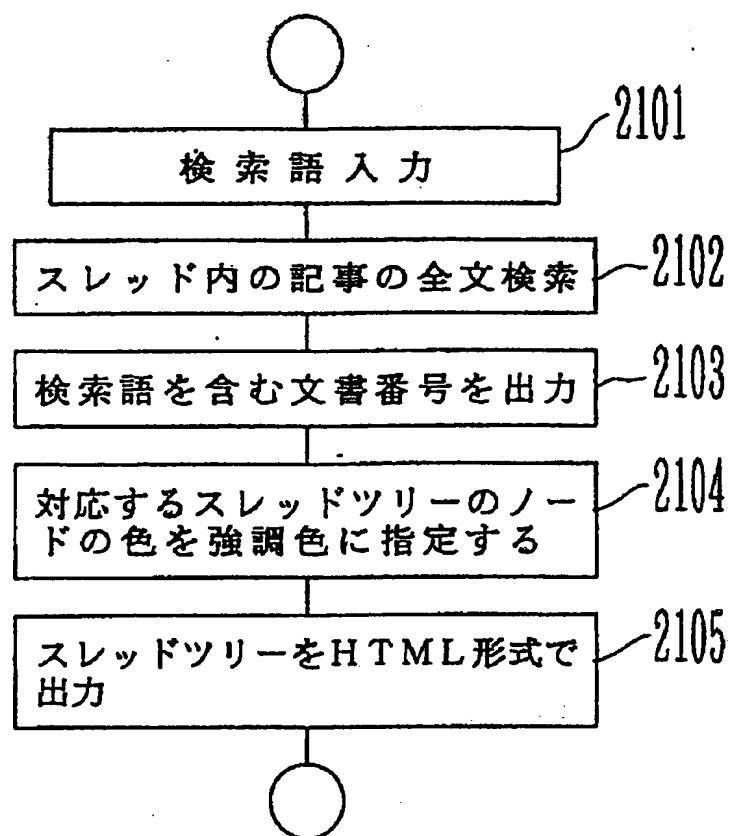
97/04/15	001, 002, 003, 004	↑	黄
97/04/21	005, 006	↑	黒
97/04/28	007, 008, 009	↑	赤
97/05/25	010	↑	青
97/05/26	011, 012, 013, 014, 015	↑	緑

(a) 作者配列の例

作者A	001, 005, 007, 008	↑	黄
作者B	011, 012, 013	↑	黒
作者C	003, 009, 010	↑	赤
作者D	004, 006	↑	青
作者E	002	↑	緑

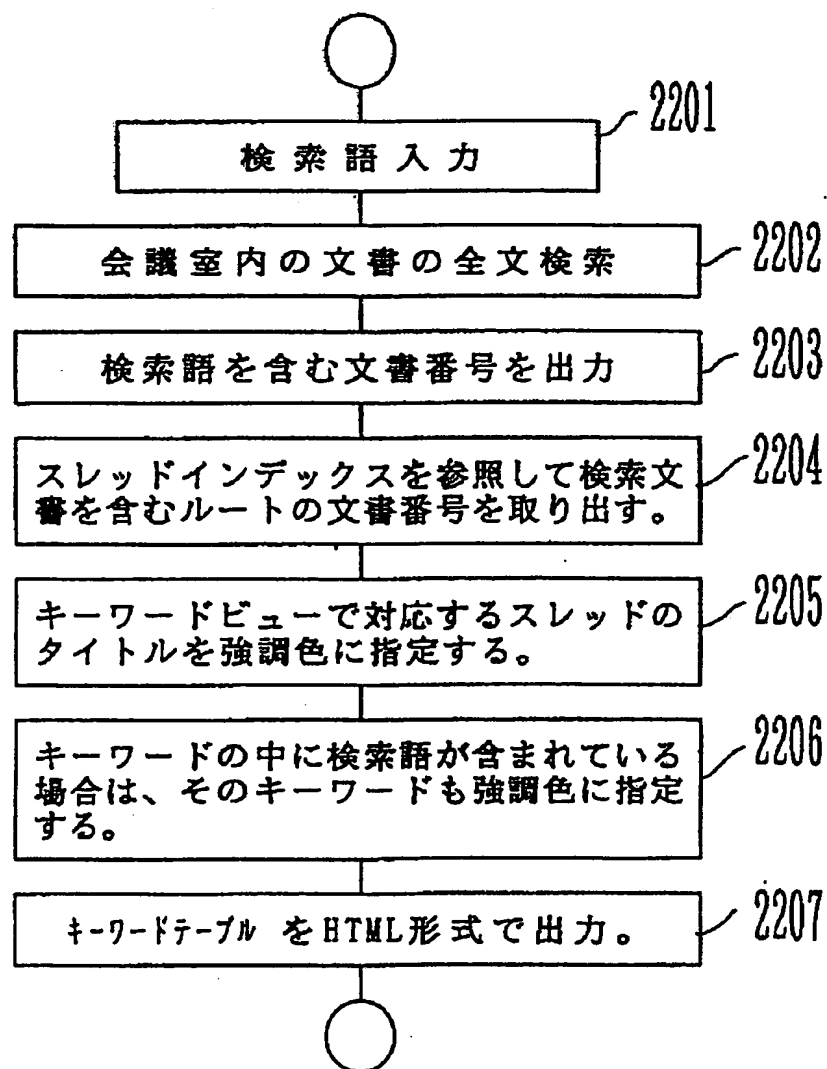
【図 2 1】

スレッドビューを使った検索結果の
強調表示の動作フローチャート



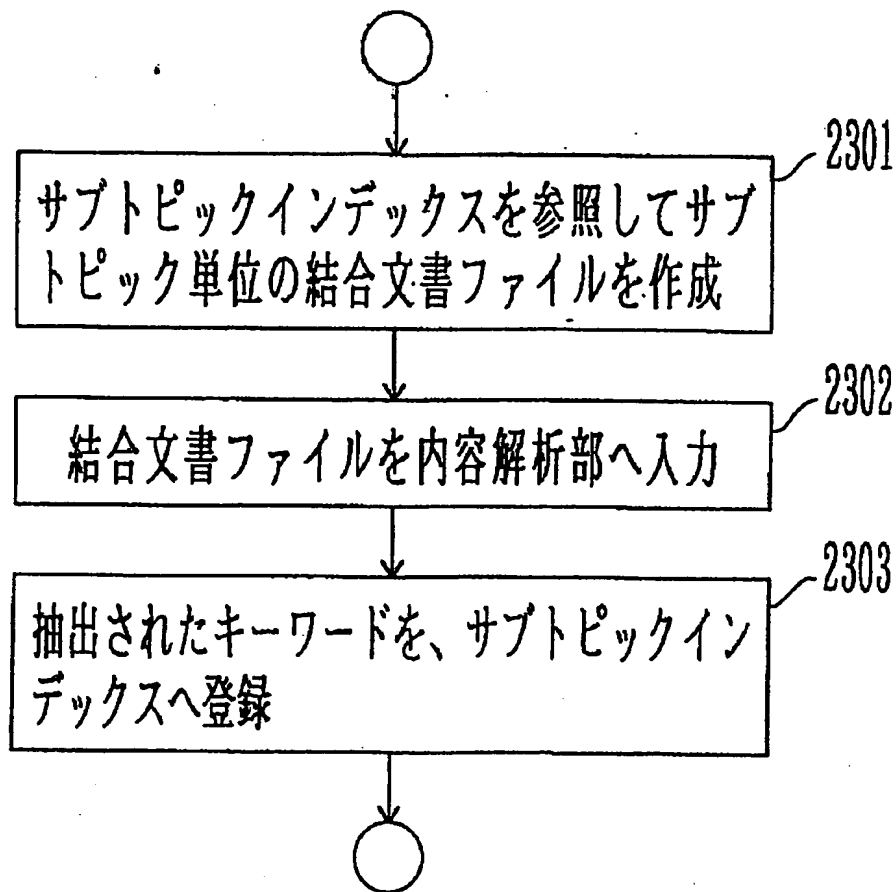
【図 2 2】

キーワードビューを使った検索結果の
強調表示の動作フローチャート



【図 2 3】

サブトピックからのキーワード抽出の
制御動作フローチャート



【図24】

サブピックインデックスの例を示す図

主文書番号	色番号	タイトル	記事番号	キーワード
∴	∴	∴	∴	∴
001	#1	(メインピク1)	001, 002, 004	eeeeee
	#2	(サブピク1)	003	eeeeee
	#3	(サブピク2)	005, 006	eeeeee
∴	∴	∴	∴	∴

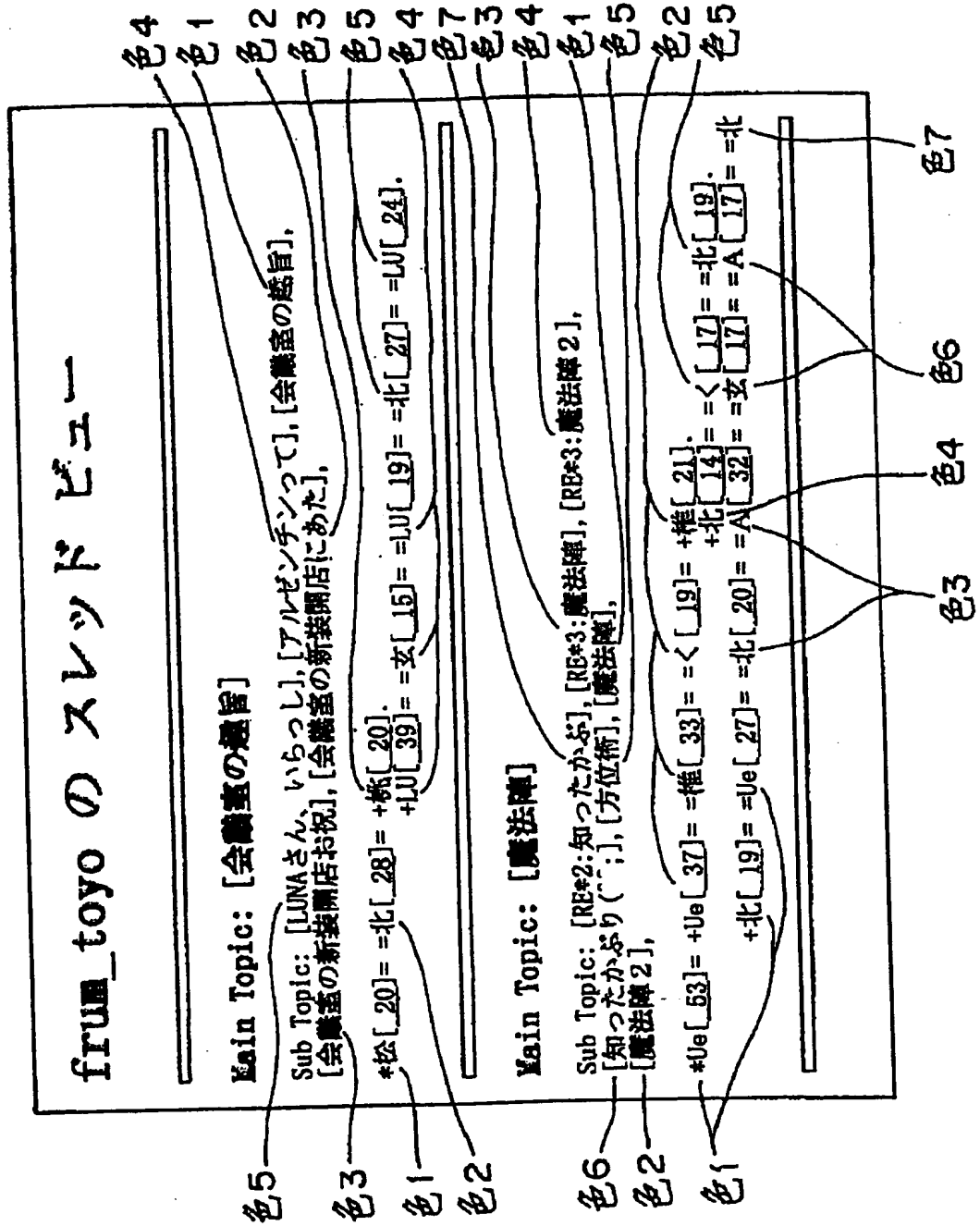
【図25】

キーワードビューの表示例を示す図

fgal_internet_QAのキーワードビュー		
[キーワードビュー] [セリダービュー]		
19	タイトル	キーワード
00001	1 はじめに	質問の内容 当合言葉 発言のタイトル Internet 関連の 質疑応答 接続形態) 用環境(ハードウェア 的確な内容の アドバイス はば/箱葉 Manager
00002	3 最低限のパソコン	最近流行のネットサーフィン インターネット デラック ・スタンダードのネットスケープ アドベーター ペンデ ・アムマシン こっちならバリ通 法律の気分 色んなサー ビス PeopleのWorldTalk MP3!
00005	3 質問: ホームページ につながらない	プロバイダ パスワードの確認 使用履歴 設定 ネット ケーブルのデフォルト ダイアルアップ接続 使用ソフト FMV&DPSP5 ユーザー名
00008	5 インターネット	プロバイダ ホームページ 料金 テレホーダイの時間開始 のコマンド www. ? 検索 534 アクロゾ95 回線速度 nifty
00014	1 皆さん町の野郎	リーヤさん推薦のプロバイダ 病半気保アバウトネットワ ーカー 明日電話 金の社員専用! ましんこ いらいらチャ ンジ せ名古屋 ROR05 この間 子会社
00015	4 ニフティのダイヤ ルアップ接続	aniftyserve つけ スクリプト SubManager or 会星人 NIFTYServe インターネット ウィザード kazusi
00016	1 プロバイダ選び	勝手気保アバウトネットワーカー オンラインサイ リー ヤさん推薦 ホームページの設定 資料請求 ソファップ 固定 料金 程度信頼 結構痛 ALLESNET
00030	9 プロバイダー 一月 宅PC間の速度	サーバーサーバー自宅PC aniftyserve プロキシ Tak ping MP5 下の方のレイヤのキャッシュ SubManager 会星人 モデムコマンド
00040	3 ニフ接続のメール	インターネットのメール aniftyserve Submanager or はば/箱葉 会星人 ニフ接続の人 NIFTYServe サービ スの中 kazusi
Windows95・サポート・Windows用のソフト まんま		

【図26】

スレッドビューの表示例を示す図



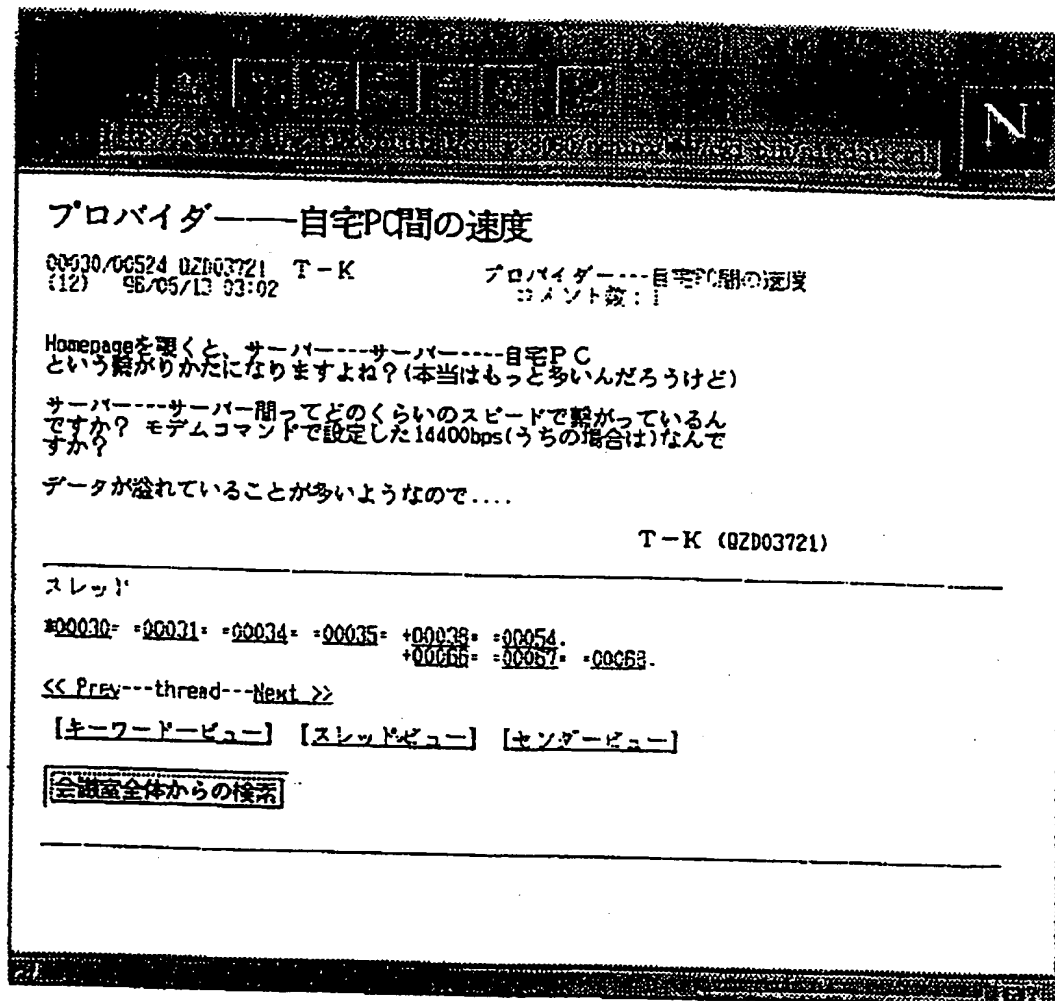
【図27】

発言者ビューの表示例を示す図

		N	
		8日:	[RE:4:SONYNE]をご存じの方おられます?
MURA	9月	11日:	[RE:NTのAdmin権限]
	30日:		[ネスケ3.0iB6でQuickTime]
	10月	1日:	[RE:2:ネスケ3.0iB6でQuickTime] [RE:NNJ.1a.0のインストール出来ず]
		2日:	[RE:ネスケ3.0iB6でQuickTime] [RE:ネスケ3.0iB6でQuickTime]
かんちゃん ★	7月	25日:	[98+MSでnewsを読み取りたい]
	27日:		[RE:98+MSでnewsを読み取りたい]
	9月	29日:	[松張子 parkについて] [RE:匿名メール]
	30日:		[RE:松張子 parkについて]
	10月	1日:	[RE:松張子 parkについて]
よしちゅ	5月	6日:	[RE:2:ニフティのダイヤルアップ接続]
	18日:		[RE:3:ニフティのダイヤルアップ接続]
	8月	5日:	[RE:2:SONYNE]をご存じの方おられます?
		6日:	[RE:3:SONYNE]をご存じの方おられます?
荒木 修	8月	3日:	[ActiveXの初歩的質問です]
	8日:		[RE:ActiveXの初歩的質問です]
	31日:		[IE3.02正式版トラブル]
	9月	2日:	[RE:IE3.02正式版トラブル]
	3日:		[RE:IE3.02正式版トラブル]
酒匂 敏弘	8日:		[ウェブサイトに繋がらない(2)]
	14日:		[RE:ウェブサイトに繋がらない(2)]
	15日:		[RE:ウェブサイトに繋がらない(2)]
	15日:		[RE:ウェブサイトに繋がらない(2)]
Mue	4日:		[RE:1:質問] さいごの印刷はどうやって]
	19日:		[RE:Huger RoadとRoad 4でのPPP接続]
	25日:		[RE:LINE SETUP WIZARDについて]
安藤貞之	7月	27日:	[トラぶってしまいました SOS!]
	28日:		[RE:トラぶってしまいました SOS!]
	30日:		[RE:トラぶってしまいました SOS!]
N-DATE	8月	18日:	[Rea Audioのプラグインについて]
	19日:		[RE:Rea Audioのプラグインについて]
	20日:		[RE:Rea Audioのプラグインについて]

【図28】

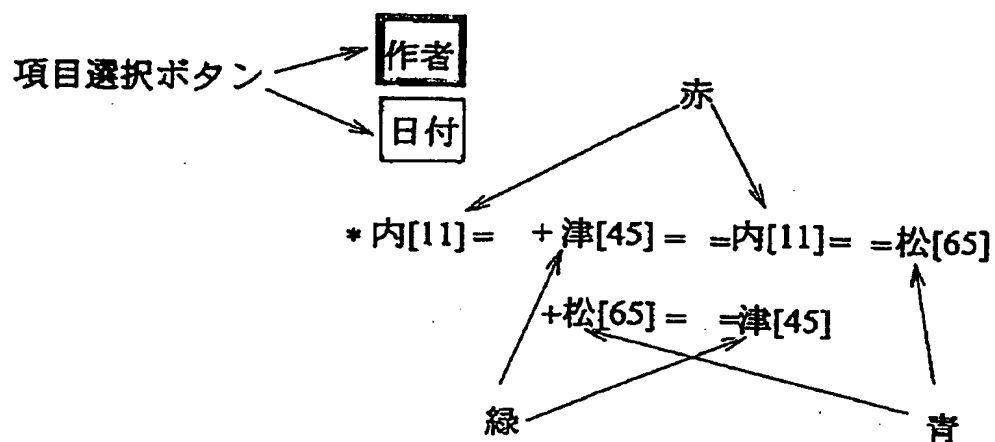
発言内容の表示例を示す図



【図 2 9】

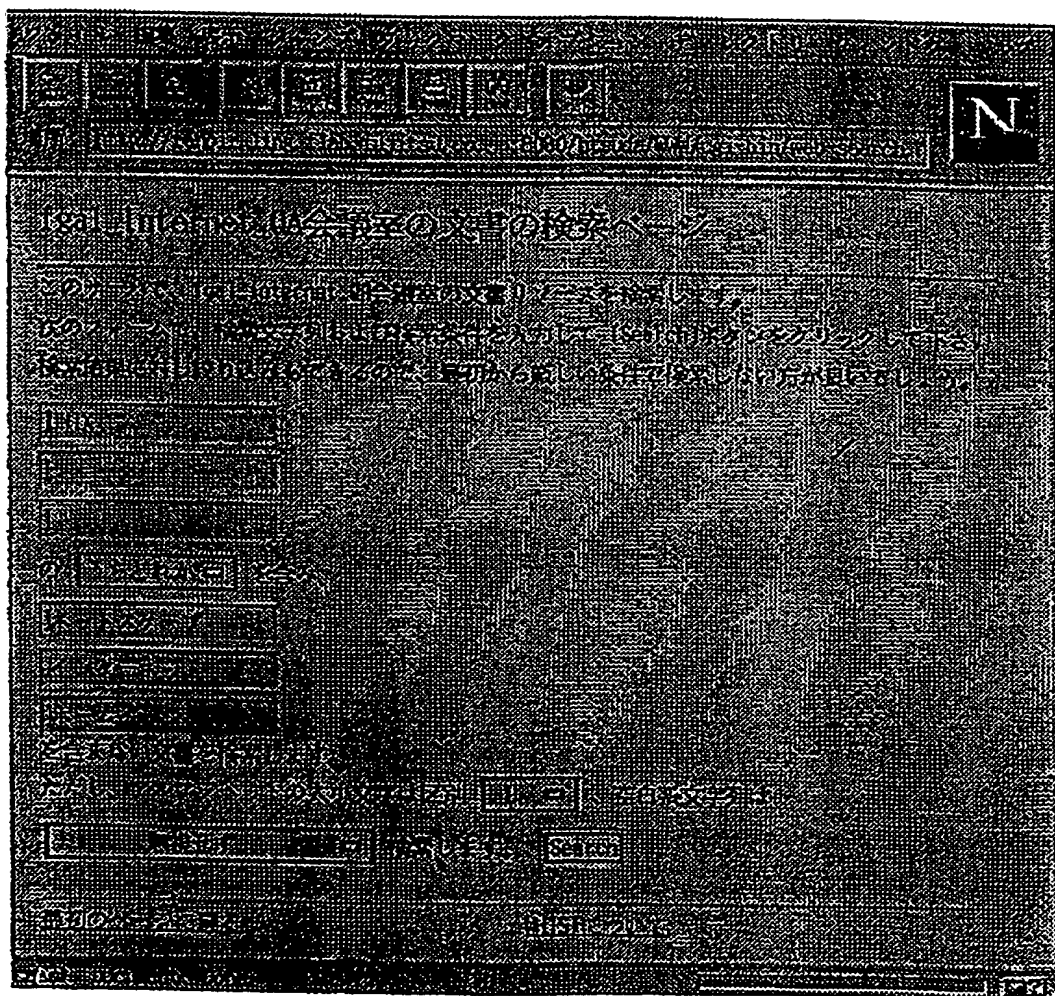
スレッドビューを用いた文書属性

「作者」の強調(色別)表示の例を示す図



【図 3 0】

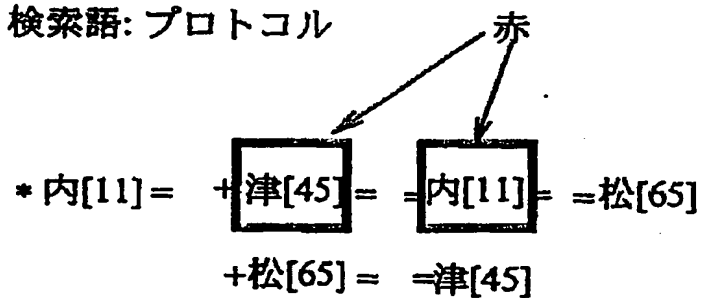
会議室内の全発言の検索(入力)表示例を
示す図



【図 3 1】

スレッドビューを用いた文字列
「プロトコル」を含むノードの強調表示の
例を示す図

検索語: プロトコル



【図 3 2】

キーワードビューを用いた文字列
「プロトコル」を含むスレッドの
強調表示の例を示す図

検索語:プロトコル

会議室名: FGL インターネット

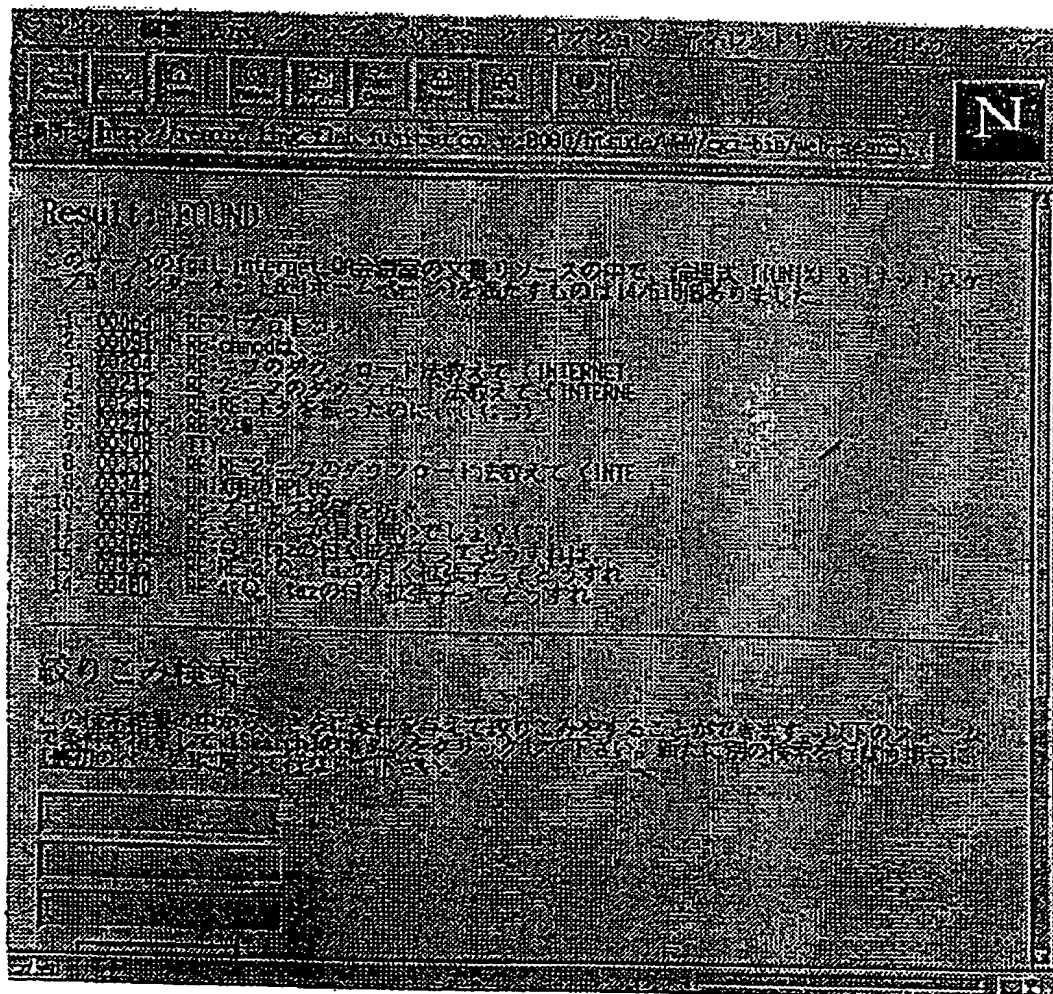
タイトル	キーワード
A 先生がケーブルを	B 大学 C 校舎 インターネットのはじまり
プロトコル	telnet ftp UNIXマシンのプロトコル
.....
Internetの規格	プロトコルの種類 http ftp
.....

赤

【圖 3 3】

会議室内の全発言の検索(結果出力)

表示例を示す図



【図 3 4】

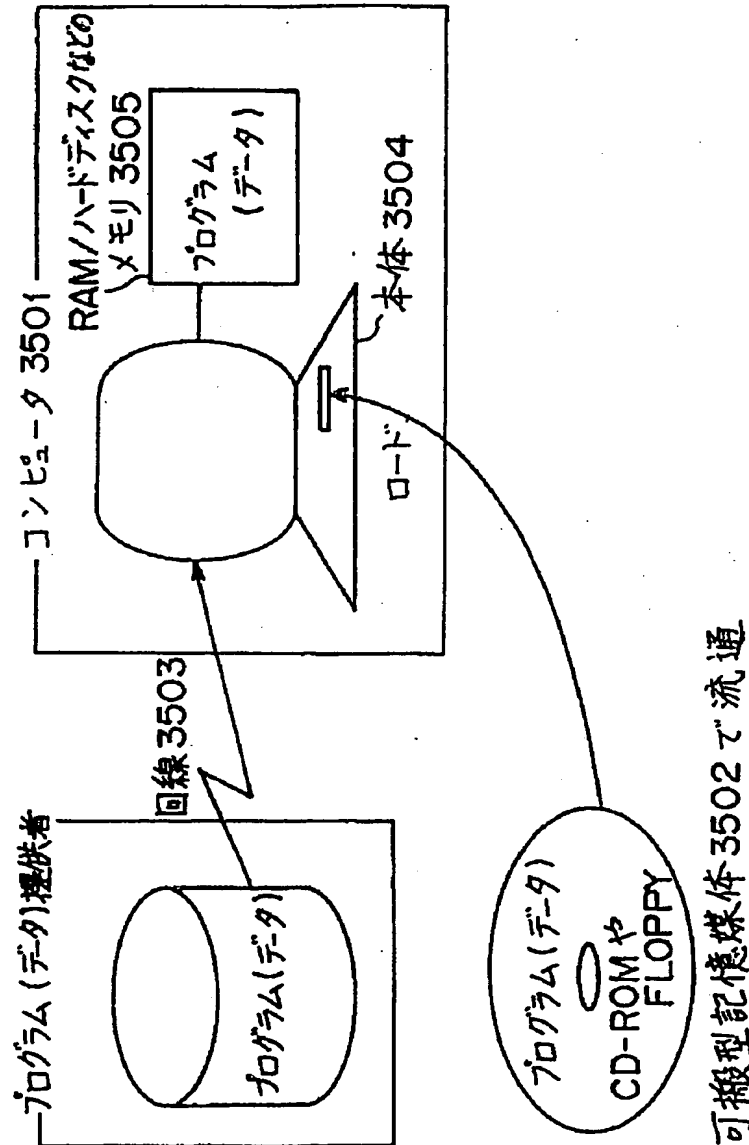
サブトピックから抽出した
キーワードの表示例を示す 図

メイントピック: 何から質問すれば

サブトピック	キーワード
A 先生がケーブルを	B 大学 C 校舎 インターネットのはじまり
プロトコル	telnet ftp UNIXマシンのプロトコル
.....

【図35】

本実施の形態を実現するプログラムが
記録された記録媒体の説明図



【書類名】 要約書

【要約】

【課題】 大量の文書が含まれる文書集合からの、必要な文書又は必要な文書が含まれているであろう文書群の検索を、支援することにある。

【解決手段】 フォーラム／会議室内の文書群は、所定のサーバコンピュータ上の文書群データベース301として、蓄積される。文書群解析装置302は、文書群データベース301内の各会議室に対応する文書集合毎に、それに含まれる文書群の解析を行う。集計装置303は、文書群解析装置302による解析結果に基づいて、メタインデックス304、スレッドインデックス305、及び索引ファイル404を生成する。表示装置306は、メタインデックス304とスレッドインデックス305を用いて、キーワードビュー、スレッドビュー、又は発言者ビューの何れかの表示形態で、文書群を表示する。

【選択図】 図3

[Document Name] Patent Application
 [Reference Number] 9701389
 [Filing Date] September 8, 1997
 [Addressee] Commissioner, Patent Office
 [International Patent Classification] G06F 17/00
 [Title of Invention] RELEVANT DOCUMENT DISPLAY APPARATUS
 AND METHOD
 [Number of Claims] 10
 [Inventor]
 [Address or Residence] c/o FUJITSU LIMITED,
 1-1, Kamikodanaka 4-chome,
 Nakahara-ku, Kawasaki-shi,
 Kanagawa
 [Name] Kanji UCHINO
 [Inventor]
 [Address or Residence] c/o FUJITSU LIMITED,
 1-1, Kamikodanaka 4-chome,
 Nakahara-ku, Kawasaki-shi,
 Kanagawa
 [Name] Hiroshi TSUDA
 [Inventor]
 [Address or Residence] c/o FUJITSU LIMITED,
 1-1, Kamikodanaka 4-chome,
 Nakahara-ku, Kawasaki-shi,
 Kanagawa
 [Name] Kunio MATSUI
 [Patent Applicant]
 [Identifying Number] 000005223
 [Name] FUJITSU LIMITED
 [Agent]
 [Identifying Number] 100074099
 [Postal Code] 102
 [Address or Residence] 3rd Fl., Nibancho Bldg.
 8-20 Nibancho, Chiyoda-ku,
 Tokyo
 [Attorney]
 [Name] Yoshiyuki OSUGA
 [Telephone Number] 03-3238-0031
 [Agent Appointed]
 [Identifying Number] 100067987
 [Postal Code] 222
 [Address or Residence] Ookurayama-Nibankan, 1418-
 305, Futoocho, Kohoku-ku,
 Yokohama-shi, Kanagawa
 [Attorney]
 [Name] Akira KUKIMOTO
 [Telephone Number] 045-545-9280
 [Fee Designation]
 [Pre-payment Register Number] 012542

[Payment Amount]	21000
[Index of Submitted Article]	
[Article Name] Specification	1
[Article Name] Drawings	1
[Article Name] Abstract	1
[General power of attorney number]	9705047
[Necessity of Proof] Yes	

[Document Name] Specification

[Title of the Invention] Relevant Document Display
Apparatus

[What is claimed is]

- 5 1. A relevant document display apparatus for
displaying a group of documents containing documents
to be cross-referenced, comprising:
- a document group analysis device for classifying
a plurality of documents forming a set of documents
10 into one or more group of documents to be cross-
referenced by analyzing a reference between the
plurality of documents forming the set of documents,
and for extracting information about classification
as document group information;
- 15 a document group keyword extraction device for
extracting a keyword contained in a document forming
part of each group of documents by referring to the
document group information; and
- a document group keyword display device for
20 displaying a title relevant to each group of documents
and a corresponding keyword extracted from the group
of documents by said document group keyword extraction
device.
- 25 2. The apparatus according to claim 1, wherein

said document group keyword device displays with enhancement a group of documents containing a document specified by a user for retrieval, or a keyword specified for retrieval.

5

3. A relevant document display apparatus for displaying a group of documents containing documents to be cross-referenced, comprising:

a document group analysis device for classifying
10 a plurality of documents forming a set of documents into one or more group of documents to be cross-referenced by analyzing a reference between the plurality of documents forming the set of documents, and for extracting information about classification
15 as document group information;

a document attribute analysis device for extracting document attribute information about an attribute of each document from a plurality of documents forming the set of documents; and

20 a document group structure display device for displaying cross-references in each group of documents in a tree structure in which the document attribute information or abbreviated information for each document forming part of the group of documents is
25 displayed as each node by referring to the document

group information and the document attribute information.

4. The apparatus according to claim 3, wherein
5 said document group structure display device
displays the cross-references in each group of
documents in a tree structure in which a combination
of an abbreviated name of an author and amount-of-
document information for each document forming part
10 of the group of documents is displayed as each node
by referring to the document group information and the
document attribute information.

5. The apparatus according to claim 3, wherein
15 said document group structure display device
further displays a plurality of topics extracted from
a document contained in each group of documents as
associated with each node forming part of the tree
structure displayed for the group of documents.

20

6. The apparatus according to claim 5, wherein
 said document group structure display device
displays each topic and a relevant node in a same
color for each group of documents.

25

7. The apparatus according to claim 3, wherein
said document group structure display device
displays with enhancement a node corresponding to a
document specified by a user for retrieval.

5

8. A relevant document display apparatus for
displaying a group of documents containing documents
to be cross-referenced, comprising:

a document group analysis device for classifying
10 a plurality of documents forming a set of documents
into one or more group of documents to be cross-
referenced by analyzing a reference between the
plurality of documents forming the set of documents,
and for extracting information about classification
15 as document group information;

a topic analysis device for further classifying
each of the classified group of documents based on
topics extracted from each document forming part of
each group of documents, and extracting information
20 about the further classification as topic
classification information;

a topic keyword extraction device for extracting
a keyword contained in a document relevant to each
topic obtained by further classifying each of the
25 groups of documents by referring to the document group

information and the topic classification information;
and

a topic keyword display device for displaying,
for each of the groups of documents and topics

- 5 obtained by further classifying the groups of
documents, a relevant title and a keyword extracted
by said topic keyword extraction device and associated
with each topic.

- 10 9. A computer-readable storage medium used to direct
a computer to perform the steps of:

classifying a plurality of documents forming a
set of documents into one or more group of documents
to be cross-referenced by analyzing a reference

- 15 between the plurality of documents forming the set of
documents;

extracting information about classification as
document group information;

- extracting a keyword contained in a document
20 forming part of each group of documents by referring
to the document group information; and

displaying a title relevant to each group of
documents and a corresponding keyword extracted from
the group of documents.

10. A computer-readable storage medium used to direct a computer to perform the steps of:

classifying a plurality of documents forming a set of documents into one or more group of documents
5 to be cross-referenced by analyzing a reference between the plurality of documents forming the set of documents;

extracting information about classification as document group information;

10 extracting document attribute information about an attribute of each document from a plurality of documents forming the set of documents; and

displaying cross-references in each group of documents in a tree structure in which the document
15 attribute information or abbreviated information for each document forming part of the group of documents is displayed as each node by referring to the document group information and the document attribute information.

20

[Detailed Explanation of the Invention]

[0001]

[Field of the invention]

The present invention relates to a technology of
25 supporting the access by users to necessary

information by properly arranging a group of cross-referenced documents and displaying the group of documents from various points of views.

[0002]

- 5 [Prior art Technology and Problems to be solved by the Invention]

In electronic conferences, electronic news, etc. established through personal computer communications or computer networks, there has been a request to
10 easily and quickly locate a necessary document in a set of sequentially accumulated documents.

[0003]

In response to the request, a conventional technology has been developed to present the users
15 with a list of titles obtained by rearranging the titles of documents in a set of documents in the order of document generation date.

[0004]

Another conventional technology displays cross-
20 references of documents by classifying a set of documents into groups of cross-referenced documents and displaying an indented title of each document in a group of cross-referenced documents. A further conventional technology displays cross-references of
25 documents by displaying the numbers of the documents

in a set of documents in a tree format.

[0005]

A further conventional technology searches entire documents in a set of documents for a specific key
5 word, and lists search results.

However, since the information displayed by the above described conventional technologies is limited, the following problems cannot be solved by the above described conventional technologies.

10

1. To retrieve a required document or a group of documents probably containing a required document from a set of various documents, the only information to be referred to is a list of titles of documents.

15 Since all titles do not correctly describe the contents of the documents, it is hard to obtain a desired retrieval result.

2. Indented titles and document numbers in a tree
20 structure are not enough to correctly inform a user of the entire structure of a group of documents and of the transition of the subjects in the group of documents.

25 3. Necessary documents cannot be accessed from

various points of view.

4. When there are a number of retrieval results,
required documents cannot be accessed unless limited
5 retrieval is performed or each of the retrieval result
lists is checked.

On the other hand, a technology of extracting a
keyword from a plurality of specific documents and
10 automatically setting a link from each document
containing a common keyword to another document is
well-known. This conventional technology enables a
plurality of relevant documents to be efficiently read
by setting the documents as cross-referenced in
15 specific documents such as published patents,
technological theses, etc.

[0006]

However, the above described conventional
technology aims at easily referencing relevant
20 documents, and cannot be used to support the retrieval
of a necessary document or a group of documents which
probably contains a necessary document.

[0007]

The present invention has been developed based
25 on the above described background to support the

retrieval of a necessary document or a group of documents which probably contains a necessary document from a set of documents containing a large volume of documents.

5 [0008]

[Means for solving the Problems]

The present invention is based on a relevant document display device for displaying a group of documents containing documents which can be cross-
10 referenced.

[0009]

The first aspect of the present invention is designed to have the following configuration.

First, a document group analysis device
15 (structure analysis unit 402) classifies a plurality of documents forming a set of documents into one or more groups of documents (thread) which can be cross-referenced by analyzing the reference among the plurality of documents forming the set of documents,
20 and extracts information relating to the classification as document group information (thread index 305).

[0010]

A document group keyword extraction device
25 (contents analysis unit 403) extracts a keyword

contained in a document forming part of each group of documents by referring to document group information.
[0011]

A document group keyword display device (display
5 device 306) displays each document with a relevant title of the document associated with a keyword extracted by the document group keyword extraction device for each group of documents.

[0012]

10 With the above described configuration according to the first aspect of the present invention, a group of documents probably containing a necessary document can be easily retrieved from a set of various documents using a keyword.

15 [0013]

According to the second aspect of the present invention, the document group keyword display device having the configuration according to the first aspect of the present invention is designed to display with
20 enhancement a group of documents containing a document specified by a user for retrieval, or a keyword specified for retrieval.

[0014]

With the configuration according to the second
25 aspect of the present invention, a group of documents

specified for retrieval can be recognized at a glance.

The third aspect of the present invention has the following configuration.

[0015]

5 The document group analysis device (structure analysis unit 402) classifies a plurality of documents forming a set of documents into one or more group of documents which can be cross-referenced by analyzing the reference among the plurality of documents forming
10 the set of documents, and then extracts information about the classification as document group information.

[0016]

 A document attribute analysis device (format
15 analysis unit 401) extracts document attribute information (meta-index 304) indicating the attribute of a document from each of a plurality of documents forming a set of documents.

[0017]

20 A document group structure display device (display device 306) displays the reference among each group of documents by referring document group information and document attribute information in a tree structure in which each document attribute
25 information or abbreviated information for each

document forming part of the group of documents is displayed as each node. To be more practical, the document group structure display device displays the reference of each group of documents among a plurality
5 of document groups in a tree structure in which an abbreviated name of an author of each document forming part of the group of documents is combined with document volume information (number of lines) to be displayed as each node.

10 [0018]

With the configuration according to the above described third aspect of the present invention, the structure of the entire group of documents can be easily recognized.

15 According to the fourth embodiment of the present invention, the document group structure display device according to the above described third aspect of the present invention further displays a plurality of topics (a main topic and a sub-topic) extracted from
20 a document contained in each group of document group by associating them with each node forming part of a tree structure displayed corresponding to the group of document. To be more practical, the document group structure display device displays for each group of
25 document each topic and each relevant node in the same

color.

[0019]

With the above described configuration according to the fourth aspect of the present invention, the transition of the subject of the entire group of documents can be easily recognized.

According to the fifth embodiment of the present invention, the document group structure display device according to the above described third aspect of the present invention is designed to display with enhancement a node corresponding to the document specified for retrieval by a user.

[0020]

With the above described configuration according to the fifth aspect of the present invention, the retrieval can be performed with the structure of the thread recognized.

The sixth aspect of the present invention has the following configuration.

[0021]

The document group analysis device (structure analysis unit 402) classifies a plurality of documents forming a set of documents into one or more groups of documents which can be cross-referenced by analyzing the reference among the plurality of documents forming

the set of documents, and extracts the information about the classification as document group information.

[0022]

5 A topic analysis device (structure analysis unit 402) further classifies each of the classified document groups based on the topic extracted from a document forming part of each group of documents, and extracts the information relating to the further
10 extraction as topic classification information (sub-topic index).

[0023]

 A topic keyword extraction device (contents analysis unit 403) extracts a keyword contained in a
15 document related to a topic for each group of documents and each of the further classified topics of each of the group of documents.

[0024]

 A topic keyword display device (display device
20 306) displays the title, of a document relating to each group of documents and each of the further detailed topics, associated with the keyword extracted for each topic by the topic keyword extraction device.

[0025]

25 With the above described configuration according

to the sixth aspect of the present invention, a precise limitation of topics can be performed using a keyword.

The present invention can be realized as a computer readable storage medium to direct a computer
5 to perform the functions similar to those performed with each of the above described configurations according to the present invention.

[0026]

[Preferred Embodiments]

10 The preferred embodiment of the present invention are described in detail below by referring to the attached drawings.

General configuration

15 FIG. 1 shows an example of a set of documents and a group of documents to be processed in the preferred embodiments of the present invention.

[0027]

In this example, there are virtual public
20 conference halls, through a computer network, referred to as forums in which users discuss for respective subjects. Each forum is divided into a plurality of virtual places referred to as message boards in which further detailed subjects are discussed about.
25 Discussions proceed with documents of messages

uploaded. Forums and message boards are defined as storage areas in a server computer, and store documents according to the above described division into virtual places. In each message board, a group of documents having a plurality of documents which can be cross-referenced forms a thread.

[0028]

A document provided by a user has, for example, a data structure shown in FIG. 2, and includes a header portion containing document attribute fields such as a document number indicating the number of the document, a date, a title, and a referenced document number to which the document refers, name of an author, etc. and a contents portion describing the body of the document.

[0029]

The following display formats are provided according to the preferred embodiment of the present invention.

20

1. Keyword view: In a set of documents forming a message board, a keyword is extracted for each thread from a group of documents forming the thread. The extracted keywords are displayed in the format shown in FIG. 25 together with the number of the documents

and the titles of the threads containing the documents.

[0030]

Through the keyword view, a user can easily find
5 a thread (group of documents) probably containing a
necessary document in a set of various documents based
on a keyword.

2. Thread view: A group of documents forming each
thread are displayed in the display format shown in
10 FIG. 26 in such a way that the references, titles,
authors, and number of lines of document can be
simultaneously recognized.

[0031]

Through the thread view, the structure of the
15 entire thread can be recognized, and the transition
of a subject can be easily understood.

3. Author view: The titles of documents are
classified based on authors, and sorted in order from
largest to smallest number of messages. For the same
20 author, the titles are sorted and displayed in the
order of date as shown in FIG. 27.

[0032]

From the viewpoint of an author and a date of a
message, a document in a set of documents (message
25 board) can be referred to.

4. Reflection of retrieval result on each view: A document relating to a user-specified retrieval keyword is displayed with enhancement in a keyword view, a thread view, etc. in the display format shown
5 in FIG. 32 or 33.

[0033]

As a result, documents can be more correctly recognized.

5. Switch of each view: A necessary document can be
10 accessed from various points of views by optionally switching the above described keyword view, thread view, and an author view.

Described below in detail are the preferred embodiments capable of displaying in the above
15 described methods according to the present invention.

[0034]

FIGs. 3 and 4 show configurations of the systems according to the preferred embodiments of the present invention.

20 The group of documents in the forum/message board are accumulated as a document group database 301 in a predetermined server computer.

[0035]

A document group analysis device 302 analyzes a
25 group of documents contained in each set of document

corresponding to each message board in the document group database 301.

A summation device 303 generates a meta-index 304, a thread index 305, and an indexed file 404 based
5 on the analysis result from the document group analysis device 302.

[0036]

A display device 306 displays a group of documents in any of the keyword view, thread view, and
10 author view formats through the meta-index 304 and the thread index 305.

[0037]

A character string retrieval device 405 retrieves each document forming part of the set of documents in
15 the document group database 301 using the indexed file 404 according to a user-specified word to be retrieved. The display device 306 reflects the retrieval result on the keyword view or the thread view and displays the result.

20 [0038]

The display device 306 displays a group of documents in any format of the keyword view, thread view, and author view using the meta-index 304 and the thread index 305.

25 [0039]

The document group analysis device 302 comprises a format analysis unit 401, a structure analysis unit 402, and a contents analysis unit 403.

The format analysis unit 401 extracts a document
5 number, title, name of author, date, and reference
document number from a header portion of each document
having the data structure shown in FIG. 2 forming part
of the set of documents in the document group database
301, computes the number of lines of the contents
10 portion of each document, and enters them in the meta-
index 304 having the data structure shown in FIG. 5
through the summation device 303.

[0040]

The structure analysis unit 402 divides each
15 document into groups of documents in thread units
based on a document number and a reference document
number extracted by the format analysis unit 401 from
each document, and generates the thread index 305
which is a reference list of documents forming each
20 thread through the summation device 303.

[0041]

FIG. 6 shows the data structure of the thread
index 305.

A list showing for each thread a root document
25 number, the number of documents, and the structure of

a thread is entered. The list is represented as follows.

(parent document number child document
5 number/sub-tree child document number/sub-tree)

The portion 'child document number/sub-tree' can further include a child list recursively.

[0042]

10 Each reference represented in each of the listings of the two threads shown in FIG. 6 is indicated to the right of the list.

The structure analysis unit 402 further divides groups of documents forming an analyzed thread into
15 groups of sub-documents having the same titles, assigns a color number to each of the groups of sub-documents, and enters color numbers assigned to the sub-document group of a document in an entry of the meta-index 304 having the data structure shown in FIG.
20 5 corresponding to the document in each sub-document.
[0043]

The contents analysis unit 403 combines a group of documents forming each thread divided by the structure analysis unit 402 in a single document file,
25 and extracts a keyword from the combined documents.

A technology of extracting a keyword from a Japanese document can be found in various well-known technologies. To extract a keyword from a Japanese document with higher precision, a character which
5 becomes a noise is removed through a pattern matching. Additionally, for example, only the predetermined number of higher order keywords are extracted.

[0044]

The keyword of each thread extracted by the
10 contents analysis unit 403 is entered in an entry of the meta-index 304 having the data structure shown in FIG. 5 corresponding to the root document of the thread through the summation device 303.

[0045]

15 The contents analysis unit 403 extracts an indexed word contained in a keyword extracted for each thread, and generates the indexed file 404 having the data structure shown in FIG. 7 through the summation device 303.

20 [0046]

The indexed file 404 is referred to by the character string retrieval device 405 as described above.

25 Detailed explanation of the document group analysis

device 302

FIG. 8 is a flowchart showing the control operations realized by the format analysis unit 401 and the structure analysis unit 402 in the document group analysis device 302 shown in FIG. 4.

[0047]

The format analysis unit 401 reads document data in line units from a new document file newly entered from the document group database 301, and extracts the document number, title, name of an author, date, and reference number from the header portion (refer to FIG. 2) of the document file (loop in steps 801, 802, 803, and back to 801).

[0048]

When the format analysis unit 401 completes extracting the data from the header portion, it generates a new entry in the meta-index 304 having the data structure shown in FIG. 5 through the summation device 303, and enters in the entry the extracted document number, title, name of an author, date, and reference document number (steps 802 and 804).

[0049]

Then, the format analysis unit 401 reads document data in line units from the contents portion (refer to FIG. 2) after the header portion in the above

described new document file, and counts the number of lines in the contents portion until the end of the file (EOF) is detected (loop in steps 805, 806, 807, and back to 805).

5 [0050]

When the end of the file is detected, the format analysis unit 401 enters the number of lines counted up to the current point through the summation device 303 in the contents portion in the entry,

10 corresponding to the document number of the new document being processed, in the meta-index 304 having the data structure shown in FIG. 5 (steps 806 and 808).

[0051]

15 In the above described line number counting process, the line recited from another document (for example, a line starting with '>') can be designed not to be counted in the line number counting process so that the number of lines actually related to the
20 message can be counted in the current document.

[0052]

Then, control is passed to the structure analysis unit 402, and the structure analysis unit 402 first enters the document number of the new document being
25 processed in the thread index 305 having the data

structure shown in FIG. 6 through the summation device 303 (step 809).

[0053]

FIG. 9 is a flowchart showing the operations
5 entered in step 809.

First, the structure analysis unit 402 determines whether or not the new document being processed is a root document of a certain thread (step 901).

Practically, when no reference document number can be
10 detected in a new document being processed in the loop in steps 801 to 803 shown in FIG. 8, the structure analysis unit 402 determines that the document is a root document.

[0054]

15 When the structure analysis unit 402 determines that a new document being processed is a root document of a certain thread, it generates a new entry in the thread index 305 through the summation device 303, and enters the document number of the new document being
20 processed in the entry as a root document number (steps 901 and 902).

[0055]

The structure analysis unit 402 initializes the number of documents in the above described entry to
25 1 after the process in step 902 (step 906), and

terminates the process in step 809 shown in FIG. 8.

When the structure analysis unit 402 determines that a new document being processed is not a root document of a certain thread, it determines whether
5 or not a list exists in an entry in the thread index 305 containing the reference document number, which is the parent document number of the list, extracted from the header portion of the new document being processed (step 901 and 903).

10 [0056]

If the structure analysis unit 402 determines that a list having as a parent document number a reference document number extracted from the header portion of a new document being processed exists in
15 the above described entry, then the structure analysis unit 402 enters the document number extracted from the header portion of the new document being processed as a child document number of the list (steps 903 and 905).

20 [0057]

If the structure analysis unit 402 determines that a list having as a parent document number a reference document number extracted from the header portion of a new document being processed does not
25 exist in the above described entry, then the structure

analysis unit 402 generates in the entry a list having the reference document number as a parent document number, and enters the document number extracted from the header portion of the new document being processed
5 as a child document number of the list (steps 903 through 905).

[0058]

After the process in step 905, the structure analysis unit 402 updates (increments by 1) the number
10 of the documents in the above described entry (step 906), and terminates the process in step 809 shown in FIG. 8.

[0059]

A practical example of the control operation
15 realized according to the flowchart of the operations shown in FIG. 9 is explained by referring to FIG. 10. FIG. 10 shows the process of generating the list of entries of the thread whose root document number is 001 in the thread index 305 shown in FIG. 6.

20 [0060]

When a root document having the document number 001 is processed, a new entry is generated in the thread index 305 by performing the processes in steps 901 and 902 in FIG. 9, the document number 001 is
25 entered as a root document number in the entry ((1)

in FIG. 10), and the number of documents of the above described entry is initialized to 1.

[0061]

Next, after a list having as a parent document
5 number the reference document number 001 extracted
from the document having the document number 002 is
generated in the entry of the root document number 001
in the thread index 305 by performing the processes
in steps 901, 903, 904, 905, and 906 shown in FIG. 9
10 when the document having the document number 002 is
processed ((2) in FIG. 10), the document number 002
is entered as a child document number of the list
(underlined portion of (3) shown in FIG. 10), and the
number of documents in the above described entry is
15 updated into 2.

[0062]

Next, after a list having as a parent document
number the reference document number 002 extracted
from the document having the document number 003 is
20 generated in the entry of the root document number 001
in the thread index 305 by performing the processes
in steps 901, 903, 904, 905, and 906 shown in FIG. 9
when the document having the document number 003 is
processed (underlined portion of (4) in FIG. 10), the
25 document number 003 is entered as a child document

number of the list (underlined portion of (5) shown in FIG. 10), and the number of documents in the above described entry is updated into 3.

[0063]

5 Next, the document number 004 is entered as a child document number of the list (underlined portion of (6) shown in FIG. 10) having as a parent document number the reference document number 003 extracted from the document having the document number 004 in
10 the entry of the root document number 001 in the thread index 305 by performing the processes in steps 901, 903, 904, 905, and 906 shown in FIG. 9 when the document having the document number 004 is processed, and the number of documents in the above described
15 entry is updated into 4.

[0064]

Next, after a list having as a parent document number the reference document number 004 extracted from the document having the document number 005 is
20 generated in the entry of the root document number 001 in the thread index 305 by performing the processes in steps 901, 903, 904, 905, and 906 shown in FIG. 9 when the document having the document number 005 is processed (underlined portion of (7) in FIG. 10), the
25 document number 005 is entered as a child document

number of the list (underlined portion of (8) shown in FIG. 10), and the number of documents in the above described entry is updated into 5.

[0065]

5 Finally, after a list having as a parent document number the reference document number 005 extracted from the document having the document number 006 is generated in the entry of the root document number 001 in the thread index 305 by performing the processes
10 in steps 901, 903, 904, 905, and 906 shown in FIG. 9 when the document having the document number 006 is processed (underlined portion of (9) in FIG. 10), the document number 006 is entered as a child document
15 number of the list (underlined portion of (10) shown in FIG. 10), and the number of documents in the above described entry is updated into 6.

[0066]

After the above described process in step 809 shown in FIG. 8, the structure analysis unit 402
20 determines the color number for the new document being processed, and performs the process of entering the color number in the entry corresponding to the document number of the above described new document in the meta-index 304 having the data structure shown
25 in FIG. 5 (step 810 shown in FIG. 8).

[0067]

FIG. 11 is a flowchart showing the entry operation in step 810. In the entry operation, a color table having the data structure shown in FIG.

5 12 is used. This table is stored in the storage device not shown in the attached drawings.

[0068]

First, the structure analysis unit 402 determines whether or not the new document being processed is a
10 root document of a certain thread (step 1101).

Practically, the structure analysis unit 402 determines that a new document being processed is a root document when no reference document number is detected in the document in the loop in steps 801
15 through 803 shown in FIG. 8.

[0069]

When the structure analysis unit 402 determines that the new document being processed is a root document of a certain thread, the entry corresponding
20 to the document number of the root document is entered in a color table having the data structure shown in FIG. 12, and the document number and the title extracted from the new document being processed (refer to step 804 shown in FIG. 8) and the initial color
25 number are entered in the entry (steps 1101, 1102, and

1103). In an example shown in FIG. 12, an entry corresponding to the color number #1 of the root document number 001 is entered, the title becomes a main topic, and only the root document number 001 is initially entered in the document number field of the entry.

[0070]

Then, the structure analysis unit 402 enters an initial color number entered in step 1103 in the entry corresponding to the document number of the above described new document in the meta-index 304 having the data structure shown in FIG. 5 (steps 1103 and 1110), thereby terminating the process in step 810 shown in FIG. 8.

15 [0071]

When the structure analysis unit 402 determines that a new document being processed is not a root document of a certain thread, it further determines whether or not the title extracted from the new document being processed (refer to step 804 shown in FIG. 8) contains a reference code such as 'Re:', etc. (steps 1101 and 1104).

[0072]

When the structure analysis unit 402 determines that the title extracted from the new document being

processed contains a reference code, the reference code is deleted from the title (steps 1104 and 1105). When the structure analysis unit 402 determines that the title extracted from the new document being
5 processed does not contain a reference code, the process in step 1105 is not performed.

[0073]

Then, the structure analysis unit 402 determines whether or not the same title that does not contain
10 the reference code and has been extracted from the new document being processed has been entered in any of the entries corresponding to the thread to which the new document being processed in the color table having the structure shown in FIG. 12 belongs (step 1106).
15 Since the thread to which the new document being processed belongs and the root document number are detected when the entry in the thread index 305 having the data structure shown in FIG. 6 is determined, the entry in the color table is determined from the root
20 document number. For example, when the new document being processed has the document number 002, an entry belonging to the root document number 001 is detected in the color table shown in FIG. 12.

[0074]

25 When the structure analysis unit 402 determines

that the same title that does not contain the reference code and has been extracted from the new document being processed has been entered in any of the entries corresponding to the thread to which the new document being processed in the color table having the structure shown in FIG. 12 belongs, the structure analysis unit 402 enters the document number of the new document being processed in the document number field of the entry (steps 1106 and 1107). For example, when the new document being processed has the document number 002, the document number 002 is entered in the document number field of the entry belonging to the root document number 001 and having an entered color number #1 in the color table shown in FIG. 12.

[0075]

Then, in the entry corresponding to the document number of the above described new document in the meta-index 304 having the data structure shown in FIG. 5, the structure analysis unit 402 enters the color number set in the entry in the color table entered in step 1107 (steps 1107 and 1110), thereby terminating the process in step 810 shown in FIG. 8.

[0076]

When the structure analysis unit 402 determines

that the same title that does not contain the reference code and has been extracted from the new document being processed has not been entered in any of the entries corresponding to the thread to which

5 the new document being processed in the color table having the structure shown in FIG. 12 belongs, a new entry corresponding to the above described thread in the color table is generated (step 1108), and a new color number in the thread, and a document number and

10 title (step 804 shown in FIG. 8) extracted from the new document being processed are entered in the generated entry (steps 1106, 1108, and 1109). For example, when the new document being processed has the document number 003, a new entry belonging to the root

15 document number 001 is generated in the color table shown in FIG. 12, and the color number #2, the title of the document having the document number 003, and the document number 003 are entered in the generated entry. This title becomes a sub-topic 1 to the main

20 topic which is the title of the root document number 001.

[0077]

Then, in the entry corresponding to the document number of the above described new document in the

25 meta-index 304 having the data structure shown in FIG.

5, the structure analysis unit 402 enters a new color number set in the new entry in the color table in step 1109 (steps 1109 and 1110), thereby terminating the process in step 810 shown in FIG. 8.

5 [0078]

The contents analysis unit 403 combines a group of documents forming each thread as a combined document file as described above by referring to the thread index 305 having the data structure shown in
10 FIG. 6, and extracts a keyword from the combined documents. As a result, the extracted keyword of each thread is entered in the entry of the meta-index 304 having the data structure shown in FIG. 5 corresponding to the root document of the thread.

15

Detailed explanation about the display device 306

As described above, the display device 306 can display a group of documents in the display format of the keyword view, the thread view, or the author view
20 using the meta-index 304 having the data structure shown in FIG. 5 and the thread index 305 having the data structure shown in FIG. 6.

[0079]

In this embodiment, if the system shown in FIG.
25 4 is designed to be connected to a Web server for

controlling the display of a home page, then a user logs in a specific message board of a specific forum by connecting a Web browser application in the terminal at hand such as a personal computer, etc. to the above described Web server, and then switches and displays the keyword view, the thread view, or the author view by clicking a predetermined GUI (graphic user interface) button using a mouse, etc.

[0080]

10 To be more practical, the display device 306 functions as, for example, a CGI (common gateway interface) application for a Web server, generates document data in an HTML (hyper text markup language) representing each of the keyword view, thread view, and author view, etc. in response to a user request received through the Web server, and transmits the generated data to the Web server. When the Web server returns the HTML document data to the user through a computer network such as the Internet, etc., the above described view is displayed on the Web browser application on the user terminal unit.

[0081]

First the operations of displaying a keyword view realized by the display device 306 are described below.

In the keyword view described above, a keyword extracted from a group of documents forming each thread is displayed in the display format shown in FIG. 25 together with the number of documents in the group of documents and the title of the thread.

[0082]

FIG. 13 is a flowchart showing the operations of displaying a keyword view performed by the display device 306.

10 First, the display device 306 reads a file of the meta-index 304 having the data structure shown in FIG. 5 (step 1301).

[0083]

Then, the display device 306 retrieves an entry
15 in which a root document is entered while reading data entry by entry from a file in the meta-index 304 (loop in steps 1301, 1302, and back to 1301). The determination of whether or not each entry is an entry in which a root document has been entered is made by the determination
20 of whether or not the value of a reference document number field of each entry is an invalid data value.

[0084]

When the display device 306 detects an entry whose root document has been entered, it converts the
25 root document number into an anchor tag whose URL

(Uniform Resource Locator), that is, united address information for an application for displaying the root document in the document group database 301 corresponding to the root document number, is

5 specified as a value of an HREF (Hypertext REference) attribute (steps 1302 and 1303).

[0085]

Then, the display device 306 obtains the number of documents (number of child documents) contained in
10 a thread by referring to the entry corresponding to the above described root document number in the thread index 305 having the data structure shown in FIG. 6 (step 1304).

[0086]

15 The display device 306 extracts a title (main topic) and a keyword from an entry whose root document has been entered in the meta-index 304 having the data structure shown in FIG. 5, and generates HTML table document data containing as one table record the above
20 described title and keyword, a root document number in the anchor tag format converted in step 1303, and a data string comprising the number of child documents obtained in step 1304 (step 1305).

[0087]

25 Next, the display device 306 generates the HTML

table document data for each thread by repeatedly performing a series of the above described processes in steps 1301 through 1305 until the end of file (EOF) is detected in the file of the meta-index 304 (steps
5 1306 and 1301).

[0088]

When the display device 306 detects the end of file (EOF) in the file of the meta-index 304 (yes as the result of the determination in step 1306), it
10 transmits the finally obtained HTML table document data to the Web server, and terminates the operations of displaying a keyword view. As a result, a keyword view is displayed in a table format shown in FIG. 25 on the Web browser application in the user terminal
15 unit.

[0089]

A user can easily find a thread probably containing a necessary document in various document sets according to a keyword of each thread in a
20 keyword view.

[0090]

The user can immediately access a root document of a desired thread by pointing to an anchor corresponding to the root document and clicking a
25 mouse button of a mouse device, etc.

[0091]

In the above described operation of displaying a keyword view, the table records of each thread can be displayed in various colors for a number of child documents so that the user can recognizes the number of messages at one glance.

[0092]

Described below are the operations of displaying a thread view realized by the display device 306.

10 In the thread view, as described above, a group of documents forming each thread are displayed in the display format shown in FIG. 26 in which a user can recognize the reference, title, name of the author, and the number of lines of a document at a glance.

15 [0093]

In FIG. 26, the reference between threads and the transition of topics are displayed by a colored tree. A node of each tree corresponds to each document, and is displayed in the following format using the leading characters (2 bytes) of the name of the author of the document and the number of lines of the document.

name of author [number of lines]

25 Symbols such as '*', '+', '=', '.', etc. are

added before and after each node. The meanings of these symbols are listed as follows.

'*' A document marked with this symbol is a root document.

5 [0094]

'+' A document referred to by the document marked with this symbol is also referred to by another document.

'=' There is a document which refers to the
10 document marked with this symbol.

[0095]

'.' There are no documents which refer to the document marked with this symbol.

15 In FIG. 26, the 'Main Topic' is followed by the title of the root document of the thread, and the 'Sub Topic' is followed by the title other than the root document in the thread. Each title is uniquely colored, and a node corresponding to the document
20 having the same title (excluding a reference code) is displayed in the same color as the title.

[0096]

Thus, the user can understand the structure of the entire thread and the transition of topics in a
25 thread.

Furthermore, each node is displayed as an anchor so that the user can immediately access the document corresponding to the node by pointing to each node and clicking a mouse button, etc.

5 [0097]

FIG. 14 is a flowchart showing the operations of displaying a thread view performed by the display device 306.

First, the display device 306 reads a list of one
10 entry (one line) corresponding to one thread, and the number of documents contained in the thread from the file of the thread index 305 having the data structure shown in FIG. 6 (step 1401). For example, in the thread index 305 having the data structure shown in
15 FIG. 6, the following list corresponding to the root document number 001 is read, and '6' is read as the number of documents.

(001 (002 003) (004 (005 006)))

20

[0098]

Next, the display device 306 restores, for example, the tree structure of the thread indicated to the right of the table shown in FIG. 6 (step 1402).
25 To represent the tree structure, the display device

306 generates the array data as shown in FIG. 15.

[0099]

Then, the display device 306 extracts an entry of the meta-index 304 having the data structure shown in FIG. 5 corresponding to a document number forming each node of the read list, and extracts from the entry the name of the author, the number of lines, the color number, and the title (step 1403). The extracted data are stored corresponding to respective nodes.

[0100]

Based on the number of the documents read in step 1401 and the color number of each node extracted in step 1403, the display device 306 determines the color of each title displayed at the head of the thread view (step 1404). This operation is realized as an operation of mapping actual colors for respective color numbers.

[0101]

Then, the display device 306 generates an HTML document for display of the title of the root document contained in a thread and other titles preceded by the 'Main Topic:' and the 'Sub Topic:' In this case, each title is sequentially read from each entry belonging to the above described thread of the color table shown

in FIG. 12 and managed by the above described structure analysis unit 402. Simultaneously, each color determined in step 1404 is computed from the number of each color sequentially read, and the display in each color is sequentially specified. Each color is specified by a color specification instruction of an HTML document (tag, etc.).

[0102]

10 Finally, the display device 306 generates an HTML document for display of a tree structure by sequentially processing data from the document number of the leftmost node forming part of the array data indicating the tree structure of the thread restored
15 in step 1402 (step 1406). In this case, as described above, the display device 306 converts the document number of each node in the tree structure into the following display data comprising the leading character (for example, 2 bytes) of the name of the
20 author of a document corresponding to the node and the number of the lines of the document based on the name of the author, the number of lines, and the color number extracted in step 1403, and then generates HTML document data for display of the display data in a
25 color corresponding to the color number of the node.

The correspondence between a color number and an actual color depends on the correspondence determined in step 1404. Furthermore, as described above, the display device 306 generates HTML document data for display of symbols such as '*', '+', '=', '.', etc. before or after the above described display data corresponding to each node based on the connection of the data. To display the tree structure in the format as is, a preformatting tag <PRE>, which is a control tag in the HTML, is used. Furthermore, the above described display data for each node is generated as an anchor tag with which the URL to the application for display of the document data in the document group database 301 corresponding to each node is specified as a value of the HREF attribute.

[0103]

Then, the display device 306 generates the view data for each thread by repeatedly performing a series of the above described processes in steps 1401 through 1406 until the end of file (EOF) is detected in the file of the thread index 305 (steps 1407 and 1401).

[0104]

When the display device 306 detects the end of file (EOF) in the file of the thread index 305 (yes as the result of the determination in step 1407), it

transmits the finally obtained HTML table document data to the Web server, and terminates the operations of displaying a thread view. As a result, a thread view is displayed in a format shown in FIG. 26 on the Web browser application in the user terminal unit.

[0105]

Described below is the operation of displaying an author view realized by the display device 306.

As described above, in the author view, the titles of documents are classified into author (author) units, and the authors are sorted in order from the highest frequency. Each author is displayed with his or her messages arranged in order of date. The above described data are displayed in the format shown in FIG. 27.

[0106]

FIG. 16 is a flowchart showing the operations of displaying the author view realized by the display device 306.

The display device 306 uses the author array data having the data structure shown in FIG. 17 to realize the author view. Then, the display device 306 initializes the author array data when the display of the author view is started (step 1601).

[0107]

Next, the display device 306 reads data of one entry from a file of the meta-index 304 having the data structure shown in FIG. 5 (step 1602).

Then, the display device 306 determines whether
5 or not the author having the name extracted from the entry is an author who is included in the author array data (step 1603).

[0108]

If the author having the name extracted from the
10 entry is an author who is not included in the author array data, then the display device 306 adds a new author item to the author array data (steps 1603 and 1604). If the author having the name extracted from the entry is an author who is included in the author
15 array data, then the display device 306 does not perform the process in step 1604.

[0109]

Next, the display device 306 enters a document
number extracted from the above described entry in the
20 corresponding author item in the author array data (step 1605).

Then, the display device 306 enters in the author
array data in author units all document numbers
entered in the meta-index 304 by repeatedly performing
25 a series of processes in steps 1602 through 1605 until

the end of file (EOF) is detected in the file of the meta-index 304.

[0110]

When the display device 306 detects the end of
5 file (EOF) in the file of the meta-index 304 (no as
the result of the determination in step 1606), the
author items in the author array data are sorted based
on the number of document numbers entered in
respective items, that is, the number of message
10 documents of each author (step 1607).

[0111]

Then, the display device 306 sorts the document
numbers in the order of date extracted from
corresponding entries in the meta-index 304 in the
15 same author item in the author array data.

[0112]

Finally, the display device 306 generates HTML
table document data for display of the name of an
author, the date and the title extracted from the
20 entry in the meta-index 304 corresponding to each
document number in an item for each author item of the
author array data obtained as a result of the sorting
operation in steps 1607 and 1608. Then, the display
device 306 transmits the generated data to the Web
25 server and terminates the operation of displaying the

author view. As a result, the author view is displayed on the Web browser application of a user terminal in the table format as shown in FIG. 27.

[0113]

5 The user can refer to a document in a set of documents (message board) from the viewpoint of an author and a date of a message in the author view.

10 The user can also refer to the message of an author in time series, and confirm the leader of authors at a glance who uploads many messages.

Other display formats of the display device 306

15 Described below are the aspects of displaying operations other than the above described operations of displaying each of the above described views realized by the display device 306.

[0114]

20 First, the operation of displaying the contents of the message realized by the display device 306 is explained below.

25 As described above, the user can immediately access a root document of each thread or a document corresponding to each node, etc. by pointing to an anchor in each thread in the keyword view or an anchor in each node in the thread view and clicking the mouse

button, etc.

[0115]

When these operations are performed by the user, the process according to the flowchart shown in FIG.

5 18 performed by the display device 306 is activated as, for example, a CGI at an instruction from the Web server. In this case, the process is provided with the information having the document number contained in the anchor tag specified by the user.

10 [0116]

As a result, after the display device 306 reads the information about the above described document number (step 1801), it reads only the document file from the document group database 301 until it reads
15 the document file containing the read document number described above in the header portion (loop in steps 1802, 1803, and back to 1802).

[0117]

When the display device 306 reads a document file
20 containing the read document number described above in the header portion (yes as a result of the determination in step 1803), it reads data line by line from the above described document file in the loop in steps 1804 through 1809 until it reads the
25 header portion of a new document. Then, it converts

the data into one line of HTML document data, and outputs the HTML document data to the Web server (step 1808).

[0118]

5 In this case, when the data in each line contains the URL for other documents, etc., the display device 306 outputs the data after converting it into an anchor tag specified as a value of an HREF attribute using the above described URL (steps 1804 and 1805).

10 [0119]

As a result, the user can jump to other resources by pointing to the anchor in the message being displayed and clicking the mouse button, etc.

When the data in each line is a comment line from
15 another document, the display device 306 adds to the data a tag for a change of color for the data (steps 1806 and 1807).

[0120]

As a result, the user can recognize a comment
20 line at a glance.

When the display device 306 completes outputting corresponding document data, it generates and outputs the HTML document in the tree structure of a thread containing the above described document, and
25 terminates the operation of displaying the contents

of the message (steps 1809 and 1810). This process can be realized as in the above described operation of displaying the thread view according to the flowchart shown in FIG. 14.

5 [0121]

As a result of the above described displaying operations, the contents of a message is displayed on the upper half of the display screen, for example, as shown in FIG. 28 on the Web browser application of the user terminal unit. The tree structure of the thread containing the document including the contents of the message is displayed on the lower half of the display screen. On the display screen, an anchor for display of a keyword view and a thread view, an anchor for use in performing retrieval, etc. can be simultaneously displayed as shown in FIG. 28.

10
15

[0122]

The function of switching the views allows the documents (messages) in a message board optionally at a request from the user, for example, from a keyword view to a thread view, the display of the contents of a message, an author view, the display of the contents of a message, the thread view, ..., etc.

20

[0123]

25 Described below is the operations of displaying

authors and dates in specified colors realized by the display device 306. FIG. 19 is a flowchart showing the above described operation.

First, the display device 306 preliminarily
5 generates the author array data in which document numbers are classified in author items as shown in FIG. 20A, and the date array data in which document numbers are classified in date items as shown in FIG. 20B based on the meta-index 304 and the thread index
10 305. The detailed explanation of the generating process is omitted here, but the process can be realized as in the process in the flowchart of FIG. 16. Different colors are assigned to author items in the author array data or date items in the date array
15 data. The assignment of color is determined from a total number of authors in each author item or a total number of dates in each date item.

[0124]

Then, the display device 306 obtains item
20 selection button information based on user specification through a Web server, and determines which has been pressed, an author button or a date selection button (refer to steps 1902 and 1904 shown in FIG. 19).

25 [0125]

If the display device 306 determines that the author button has been pressed, the display device 306 generates and outputs the display data for a thread tree by referring to the author array data shown in FIG. 20A (steps 1902 and 1903). This process is realized as in the above described process shown in the flowchart of FIG. 14. In this case, each node of the tree is displayed in a color assigned to the author item in the author array data corresponding to the name of an author corresponding to the node.

[0126]

When the display device 306 determines that the date selection button has been pressed, it generates and outputs thread tree display data by referring to the date array data shown in FIG. 20B (steps 1904 and 1905). This process is also realized as in the above described process according to the flowchart of FIG. 14. In this case, each node of the tree is displayed in the color assigned to the date item in the date array data corresponding to the date corresponding to the node.

[0127]

As a result of the above described displaying operations, a tree of a thread uniquely colored for each author is displayed, for example, as shown in

FIG. 29 on the Web browser application of the user terminal when the user presses the author button. Therefore, the user can recognize at a glance the documents written by the same author.

5 [0128]

Described below is the operation of enhancing the display of the retrieval result using a thread view realized by the display device 306. FIG. 21 is a flowchart showing the operation.

10 First, the display device 306 generates and outputs the HTML document data for display of a post-retrieval input form screen (step 2101). As a result, a retrieval input form screen as shown in FIG. 30 is displayed on the Web browser application of the user
15 terminal. The user specifies the execution of retrieval by inputting a word to be retrieved in the retrieval input form.

[0129]

The word to be retrieved is input in the
20 retrieval input form and passed to the character string retrieval device 405 (FIG. 4) through the Web server. The character string retrieval device 405 searches the entire sentences in each document forming part of a thread specified in the document group
25 database 301 using the indexed file 404 for the user-

specified word to be retrieved, and outputs the document number containing the word to be retrieved (steps 2102 and 2103).

[0130]

5 When the display device 306 receives a document number containing the above described word to be retrieved, it displays the HTML document data in the tree structure of the thread containing the document number in the process according to the flowchart of
10 FIG. 14. In this case, the display device 306 specifies the color of the node containing the above described document number as an enhanced color (steps 2104 and 2105).

[0131]

15 As a result, for example, a thread view as shown in FIG. 31 is displayed with enhancement on the Web browser application of the user terminal based on the retrieval result. Thus, the user can recognize the structure of a thread and performs the retrieval.

20 [0132]

Described below is the operation of displaying with enhancement a retrieval result using a keyword view realized by the display device 306. FIG. 22 is a flowchart showing the operations.

25 First, the display device 306 generates and

outputs the HTML document data for display of the post-retrieval input form screen as in step 2101 shown in FIG. 21 (step 2201). The user specifies performing retrieval by inputting a word to be retrieved to a
5 retrieval input form.

[0133]

The word to be retrieved and input to the above described retrieval input form is passed to the character string retrieval device 405 (FIG. 4) through
10 the Web server. The character string retrieval device 405 retrieves all sentences in each document forming part of the message board specified in the document group database 301 using the indexed file 404 by the user specifying a word to be retrieved, and outputs
15 the document number containing the word to be retrieved (steps 2202 and 2203).

[0134]

When the display device 306 receives the document number containing the above described word to be
20 retrieved, it first refers to the thread index 305 having the data structure shown in FIG. 6 and extracts the root document number corresponding to the entry containing the above described document number (step 2204).

25 [0135]

Then, the display device 306 displays the HTML document data for display of the keyword view about a specified message board in the process as shown in the flowchart of FIG. 13. In this case, the display
5 device 306 specifies as an enhanced color the color of the title or the entire display area of the thread corresponding to the root document number extracted in step 2204, and also specifies the color of a
10 keyword as an enhanced color when the displayed keyword contains a word to be retrieved (steps 2205, 2206, and 2207).

[0136]

As a result, for example, a keyword view as shown in FIG. 32 is displayed with enhancement on the Web
15 browser application of the user terminal based on the retrieval result. Thus, the user can recognize a thread containing a word to be retrieved at a glance.

[0137]

The display device 306 can be designed to display
20 the document number as a retrieval result and the corresponding title as a listing as shown in FIG. 33.

[0138]

Described finally is the operation of displaying a keyword view for each sub-topic realized by the
25 display device 306.

In the above described keyword view, a keyword is displayed for each thread. On the other hand, in the keyword view for each sub-topic, a keyword can be extracted and displayed for each sub-topic in a
5 thread.

[0139]

In this operation, the display device 306 uses a sub-topic index having the data structure shown in FIG. 24. The sub-topic index has the data structure
10 in which a keyword field is added to the data structure of the color table shown in FIG. 12.

[0140]

A sub-topic index is used to actually replace the above described color table. Therefore, the contents
15 of the field other than a keyword field in the sub-topic index are preliminarily entered in the above described process in step 810 shown in FIG. 8 performed by the structure analysis unit 402. In this case, there can be an entry obtained corresponding to
20 each of the main topic indicating the title of a root document contained in each root document number (thread) and the sub-topics indicating the titles of other documents in the sub-topic index which is a color table as described above about the process in
25 step 810 shown in FIG. 8. The display device 306 uses

the contents of the entries.

[0141]

FIG. 23 is a flowchart showing the operations of controlling the extraction of a keyword from the sub-
5 topic realized by the display device 306.

First, the display device 306 puts groups of documents in each of the combined document files in main topic units and sub-topic units based on the document number entered in each entry contained in
10 each thread in the sub-topic index, and inputs a resultant combined document file in the contents analysis unit 403 (FIG. 4) (step 2302).

[0142]

The contents analysis unit 403 extracts a keyword
15 for each combined document file, and returns a result to the display device 306. The display device 306 enters the keyword for each combined document file returned from the contents analysis unit 403 in the keyword field of the entry corresponding to the above
20 described combined document file in the sub-topic index (step 2303).

[0143]

As described above, a keyword is extracted from a group of documents belonging to each thread in main
25 topic and sub-topic units.

Afterwards, the display device 306 generates and outputs the HTML document data for display of a title and a keyword belonging to each title for a thread specified by a user in main topic and sub-topic units
5 for each thread based on the contents of the sub-topic index.

[0144]

As a result, a keyword view for each sub-topic is displayed in the format as shows in FIG. 34 on the
10 Web browser application of the user terminal. Thus, the user can precisely limit the topics using a keyword.

Supplement to storage medium storing a program for realizing a preferred embodiment of the present
15 invention

The present invention can also be designed as a computer-readable storage medium used to direct a computer to perform the functions realized by each of the configurations according to the preferred
20 embodiments of the present invention.

[0145]

In this case, for example, a portable storage medium 3502 such as a floppy disk, a CD-ROM disk, an optical disk, a removable hard disk, etc., and a
25 program for realizing each step according to the

preferred embodiment of the present invention through a network line 3503 are loaded onto memory (RAM, a hard disk, etc.) 3505 in a body 3504 of a computer 3501 for execution as shown in FIG. 35.

5 [0146]

[Effect of the Invention]

According to the first aspect of the present invention, a group of documents probably containing a required document can be precisely and easily detected
10 from a set of various documents.

[0147]

According to the second aspect of the present invention, a group of documents to be referred to can be obtained at a glance.

15 According to the third aspect of the present invention, the entire structure of a group of documents can be easily obtained.

[0148]

20 According to the fourth aspect of the present invention, the transition of the subjects in the group of documents can be easily obtained.

According to the fifth aspect of the present invention, retrieval can be carried out while obtaining the structure of thread.

25 [0149]

According to the sixth aspect of the present invention, a topic can be more precisely selected using a keyword.

Thus, according to the present invention, a group
5 of documents cannot be accessed from various points of view.

[Brief Description of the Drawings]

FIG. 1 shows an example of group of documents which can be cross-referenced;

10 FIG. 2 shows an example of the data structure of a document;

FIG. 3 shows the configuration (1) of the system according to a preferred embodiment of the present invention;

15 FIG. 4 shows the configuration (2) of the system according to a preferred embodiment of the present invention;

FIG. 5 shows the structure of meta-indexed data;

20 FIG. 6 shows the structure of thread-indexed data;

FIG. 7 shows the data structure of an indexed file;

FIG. 8 is a flowchart showing the operations of the format analysis unit and the structure analysis
25 unit;

FIG. 9 is a flowchart showing the operations when an entry of a document number is made in the thread index of a document number;

FIG. 10 illustrates the operations when an entry
5 of a document number is made in the thread index of a document number;

FIG. 11 is a flowchart showing the operations of entering color numbers;

FIG. 12 shows an example of a color table;

10 FIG. 13 is a flowchart showing the operation of a keyword view;

FIG. 14 is a flowchart showing the operation of a thread view;

15 FIG. 15 is an example of a control array for a thread view;

FIG. 16 is a flowchart showing the operations of an author view;

FIG. 17 shows an example of a control array for an author view;

20 FIG. 18 is a flowchart showing the operations for displaying the contents of a message;

FIG. 19 is a flowchart showing the operations of displaying authors and dates in respective colors.

25 FIGS. 20A and 20B show examples of display arrays for authors and dates in respective colors;

FIG. 21 is a flowchart showing the operations of a highlighted display of a retrieval result using thread views;

FIG. 22 is a flowchart showing the operations of
5 a highlighted display of a retrieval result using keyword views;

FIG. 23 is a flowchart showing the operations of controlling the extraction of a keyword from a sub-topic;

10 FIG. 24 shows an example of a sub-topic index;

FIG. 25 shows an example of displaying a keyword view;

FIG. 26 shows an example of displaying a thread view;

15 FIG. 27 shows an example of displaying an author view;

FIG. 28 shows an example of displaying the contents of a message;

FIG. 29 shows an example of displaying with
20 highlight (in different colors) a document attribute 'author' using a thread view;

FIG. 30 shows an example of displaying retrieval (input) of all messages in a message board;

FIG. 31 shows an example of displaying with
25 highlight a node containing a character string

'protocol' using a thread view;

FIG. 32 shows an example of displaying with highlight a thread containing a character string 'protocol' using a keyword view;

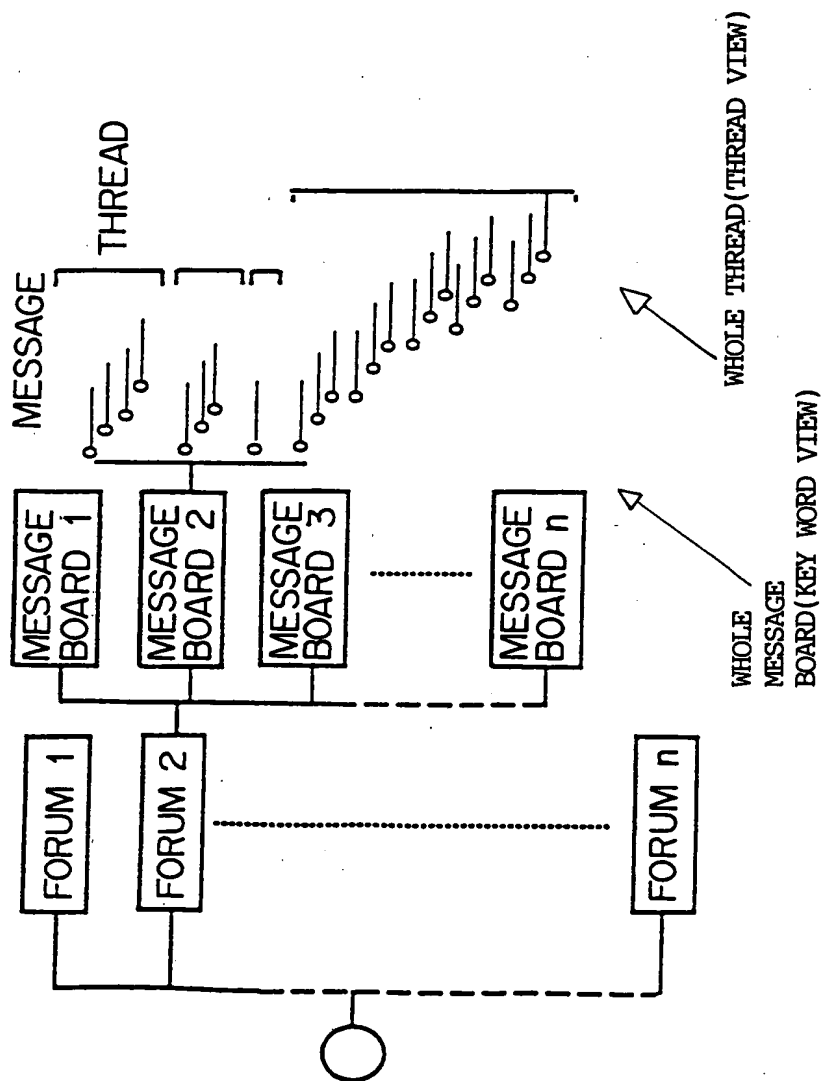
5 FIG. 33 shows an example of displaying retrieval (output of result) of all messages in a message board;

FIG. 34 shows an example of displaying a keyword extracted from a sub-topic; and

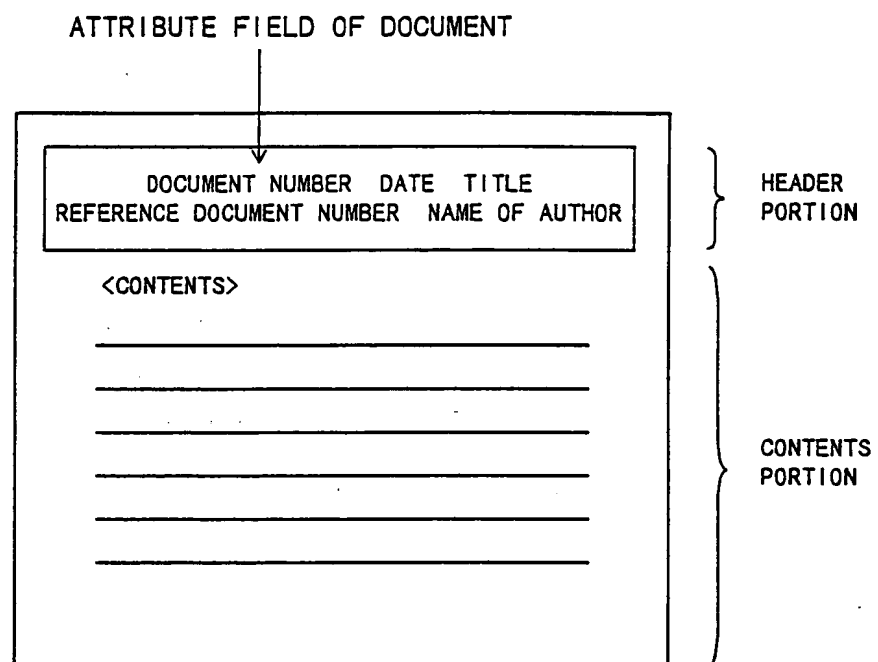
FIG. 35 shows a storage medium storing a program
10 for realizing a preferred embodiment of the present invention.

[Explanation of the Codes]

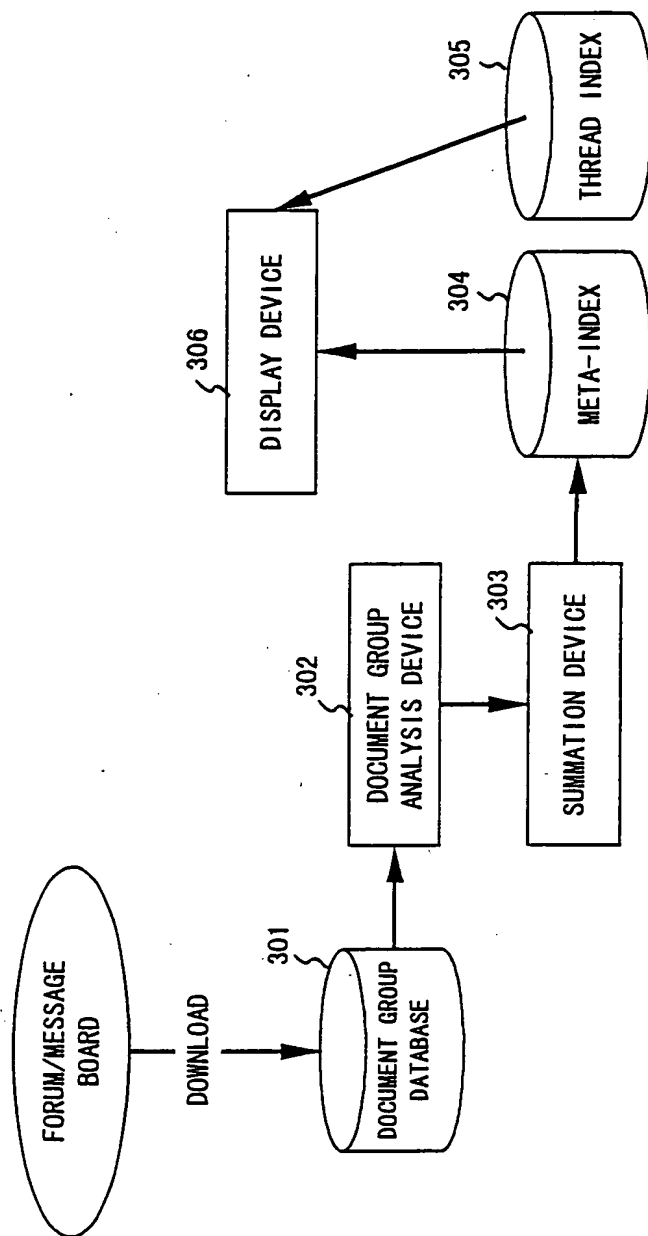
	301	Document data group
	302	Document group analysis device
15	303	Summation device
	304	Meta-index
	305	Thread index
	306	Display device
	401	Format analysis unit
20	402	Structure analysis unit
	403	Contents analysis unit
	404	Indexed file
	405	Character string retrieval device



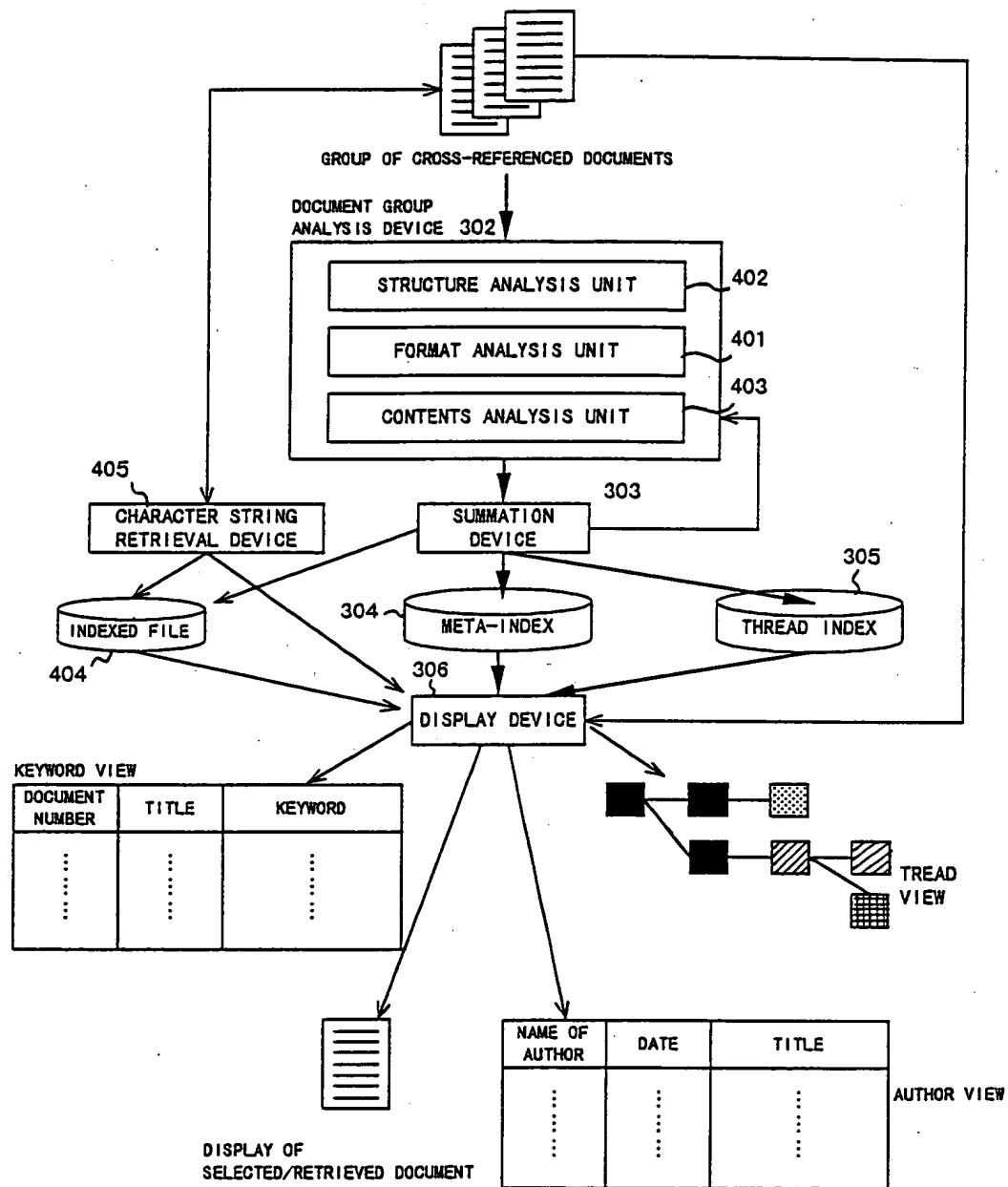
[FIG. 1] EXAMPLE OF GROUP OF DOCUMENTS WHICH CAN BE CROSS-REFERENCED



[FIG. 2] EXAMPLE OF THE DATA STRUCTURE OF A DOCUMENT



[FIG. 3] CONFIGURATION (1) OF THE SYSTEM ACCORDING TO
A PREFERRED EMBODIMENT OF THE PRESENT INVENTION



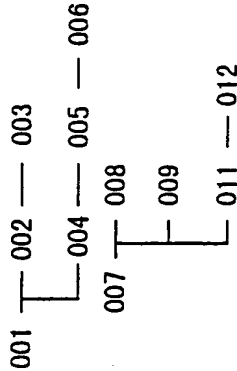
[FIG. 4] CONFIGURATION (2) OF THE SYSTEM ACCORDING TO A PREFERRED EMBODIMENT OF THE PRESENT INVENTION

DOCUMENT NUMBER	REFERENCE DOCUMENT NUMBER	TITLE	NAME OF AUTHOR	KEYWORD	NUMBER OF LINES	COLOR NUMBER	DATE
:	:	:	:	:	:	:	:
004	001	TELL ME	UCHINO	-----	12	4	96/10/05
005	004	Re: TELL ME	TSUDA	-----	45	4	96/10/08
006	005	APPLICATION OF Java	UCHINO	-----	80	6	96/10/15
*007	-----	DEFECTIVE HD?	MATSUI	SunOS, HD, 64MB, FORMAT	12	1	96/10/18
:	:	:	:	:	:	:	:

* INDICATES ROOT DOCUMENT NUMBER

[FIG. 5] STRUCTURE OF META-INDEXED DATA

ROOT DOCUMENT NUMBER	NUMBER OF DOCUMENTS	STRUCTURE OF THREAD
:		
001	6	(001 (002 003) (004 (005 006)))
007	5	(007 008 009 (011 012))
⋮ ⋮ ⋮		



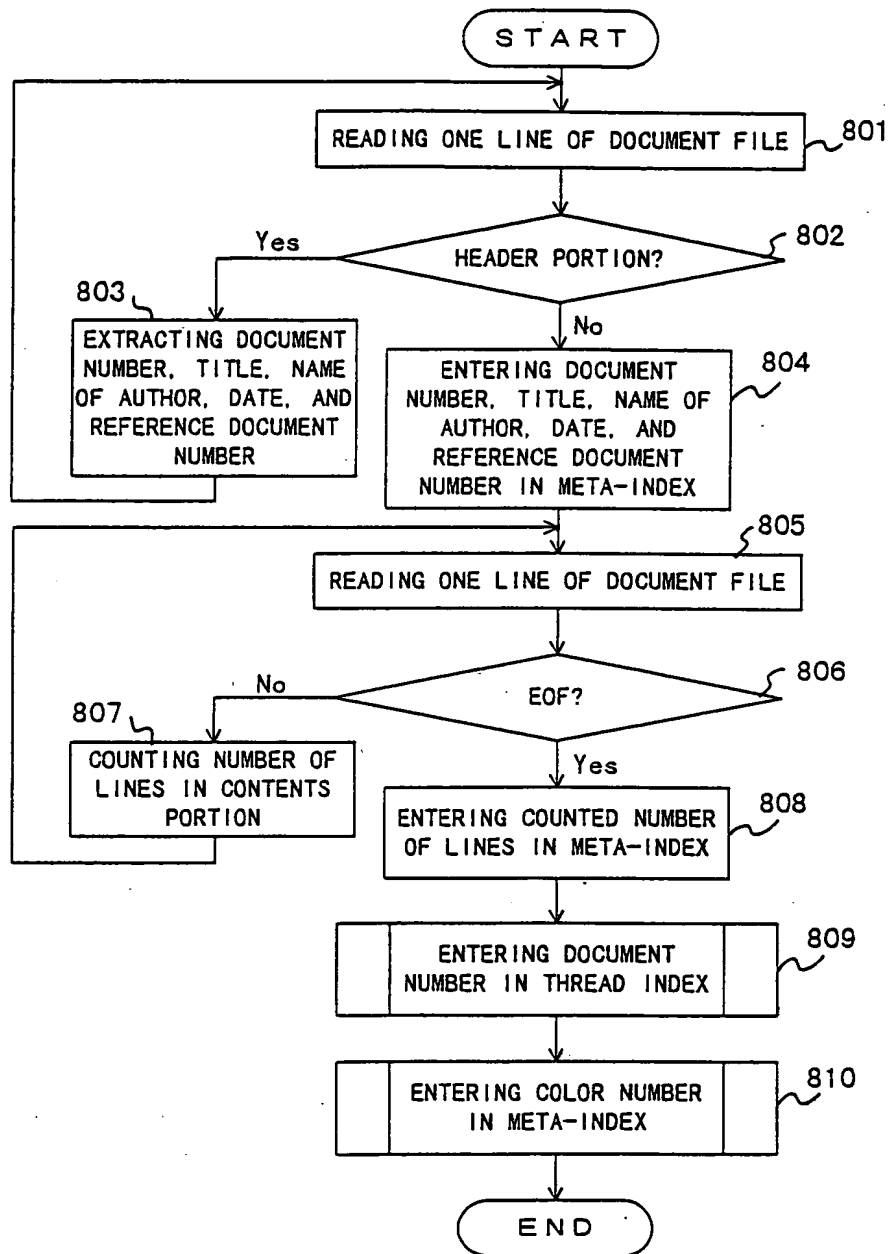
MEANING OF LISTING: (PARENT DOCUMENT NUMBER CHILD DOCUMENT NUMBER/SUB-TREE CHILD DOCUMENT NUMBER/SUB-TREE)

[FIG. 6] STRUCTURE OF THREAD-INDEXED DATA

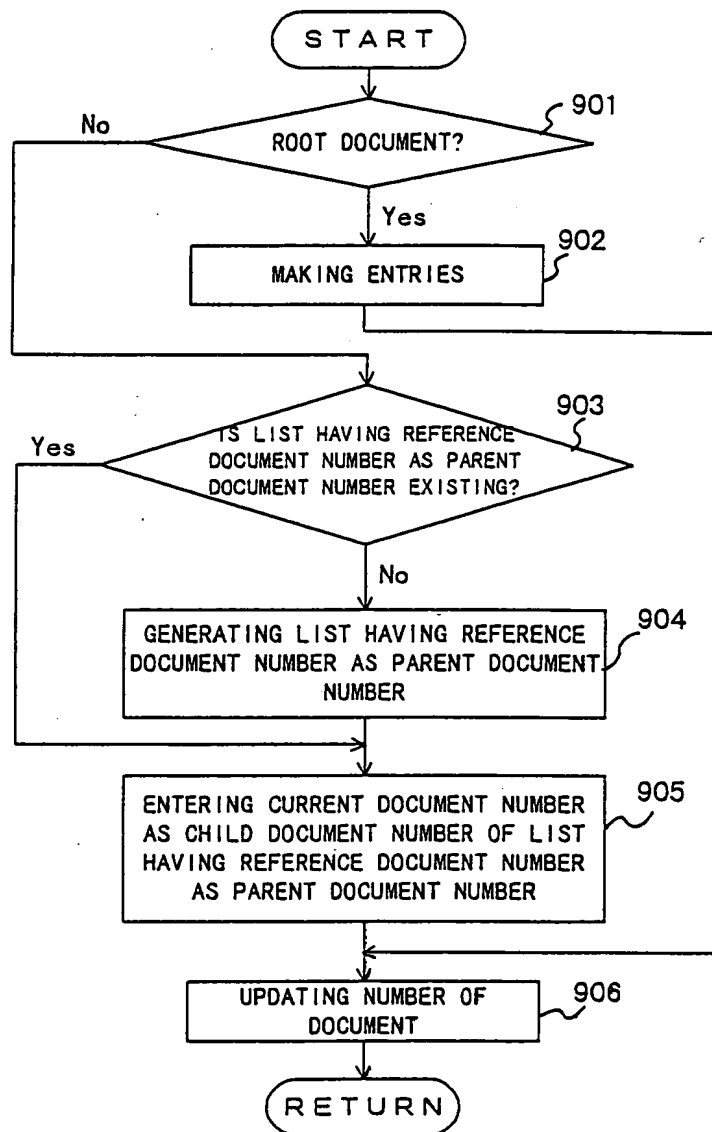
INDEXED WORD	KEYWORD CONTAINING INDEXED WORD	THREAD NUMBER*
INTERNET	·STARTING INTERNET	00496
	·CONNECTION TO INTERNET	00218
CONTROL	·CONTROL PANEL	00518
DOMAIN	·DOMAIN NAME SERVER	00128
	·DOMAIN IN NATION	00476
⋮	⋮	⋮

* THREAD NUMBER MEANS ROOT DOCUMENT NUMBER OF THREAD.

[FIG. 7] DATA STRUCTURE OF AN INDEXED FILE

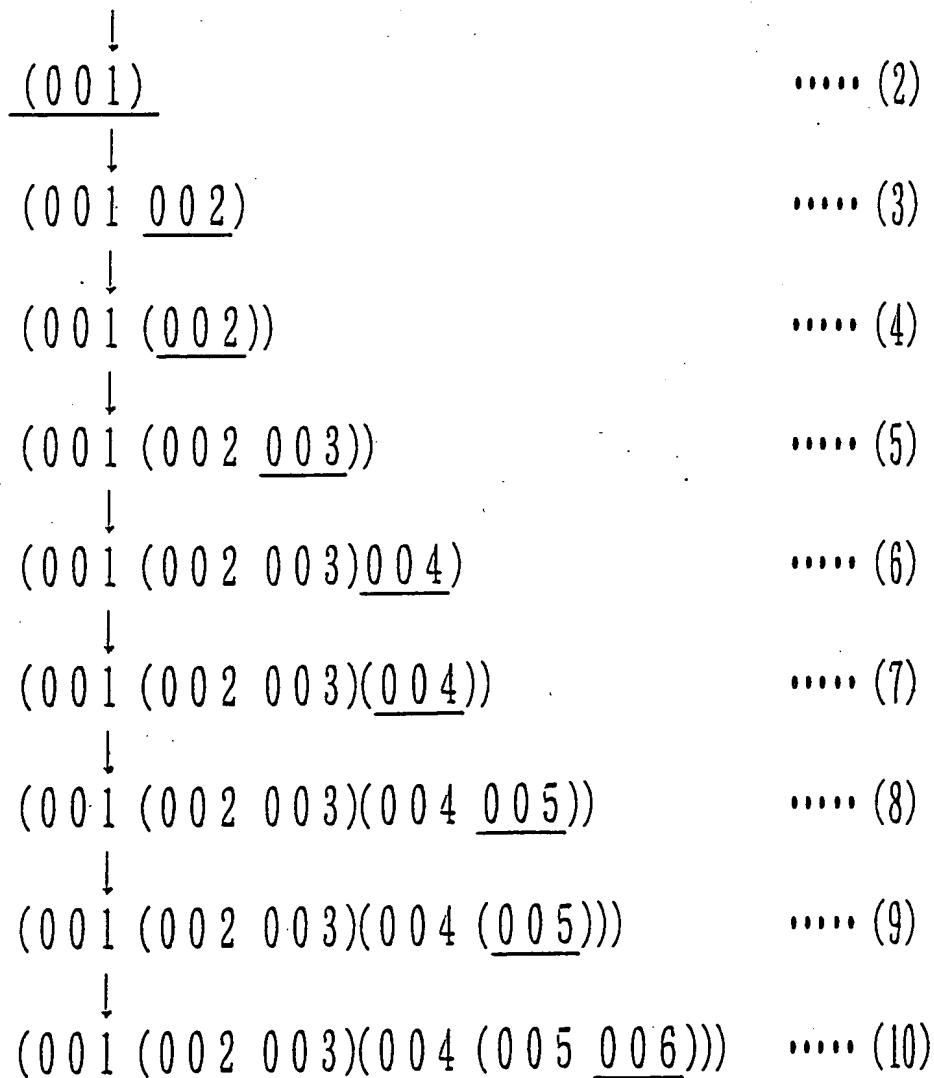


[FIG. 8] FLOWCHART SHOWING THE OPERATIONS OF THE FORMAT ANALYSIS UNIT AND THE STRUCTURE ANALYSIS UNIT

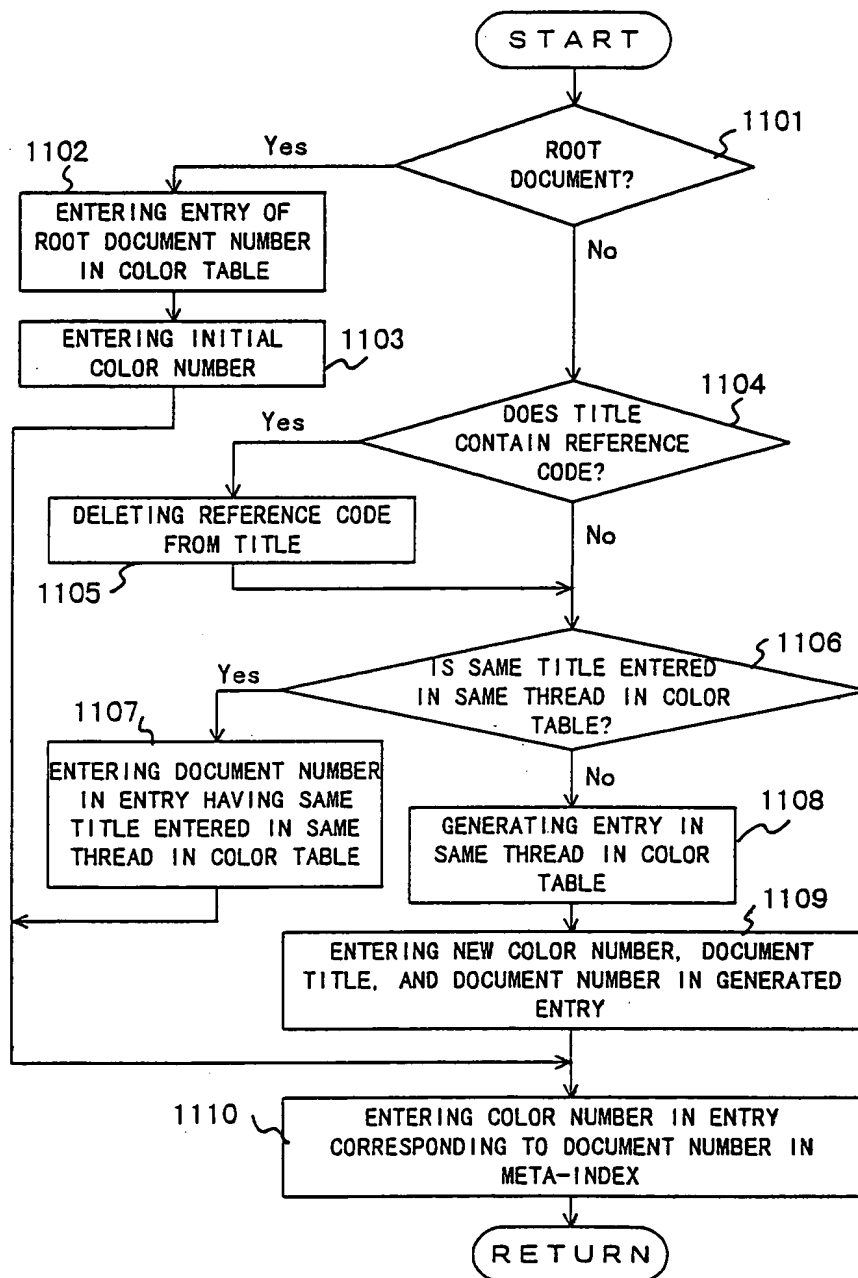


[FIG. 9] FLOWCHART SHOWING THE OPERATIONS WHEN AN ENTRY OF A DOCUMENT NUMBER IS MADE IN THE THREAD INDEX

MAKING ENTRIES



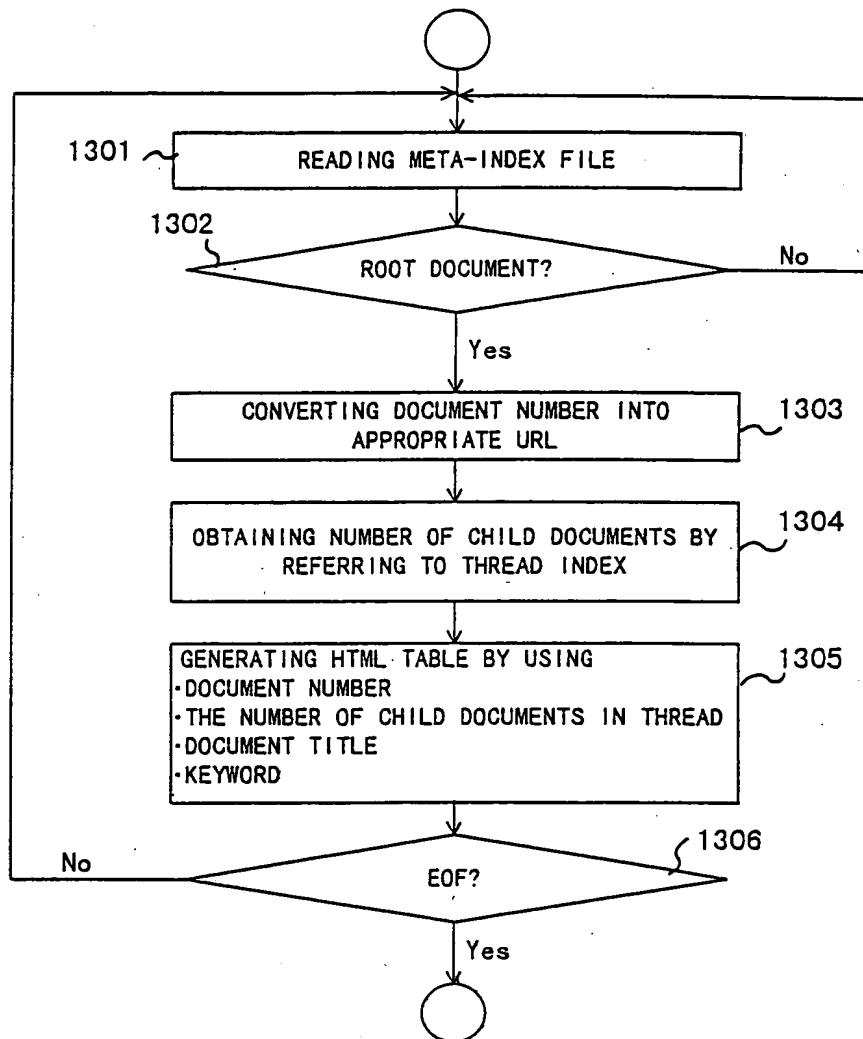
[FIG. 10] OPERATIONS WHEN AN ENTRY OF A DOCUMENT NUMBER
IS MADE IN THE THREAD INDEX



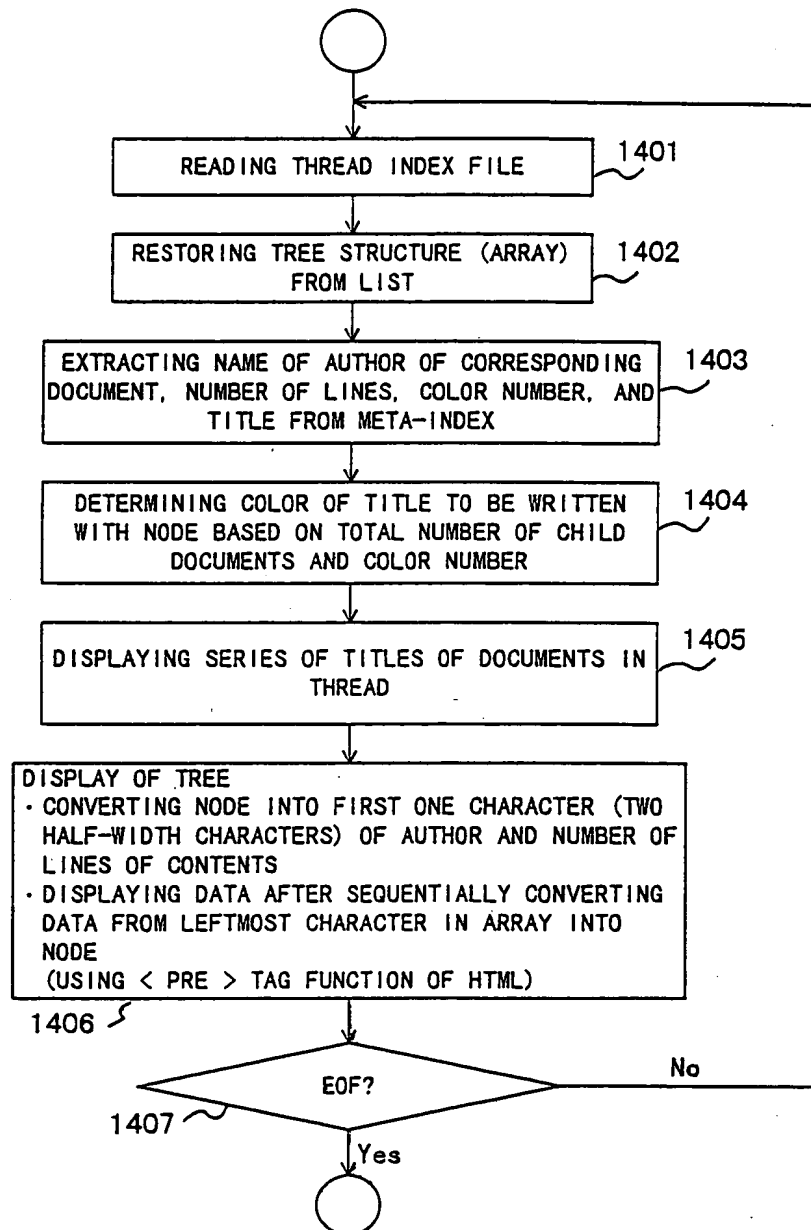
[FIG. 11] FLOWCHART SHOWING THE OPERATIONS OF ENTERING COLOR NUMBERS

ROOT DOCUMENT NUMBER	COLOR NUMBER	TITLE	DOCUMENT NUMBER
⋮	⋮	⋮	⋮
001	#1	(MAIN TOPIC)	001, 002, 004
	#2	(SUB-TOPIC 1)	003
	#3	(SUB-TOPIC 2)	005, 006
⋮	⋮	⋮	⋮

[FIG. 12] EXAMPLE OF A COLOR TABLE



[FIG. 13] FLOWCHART SHOWING THE OPERATION OF A KEYWORD VIEW



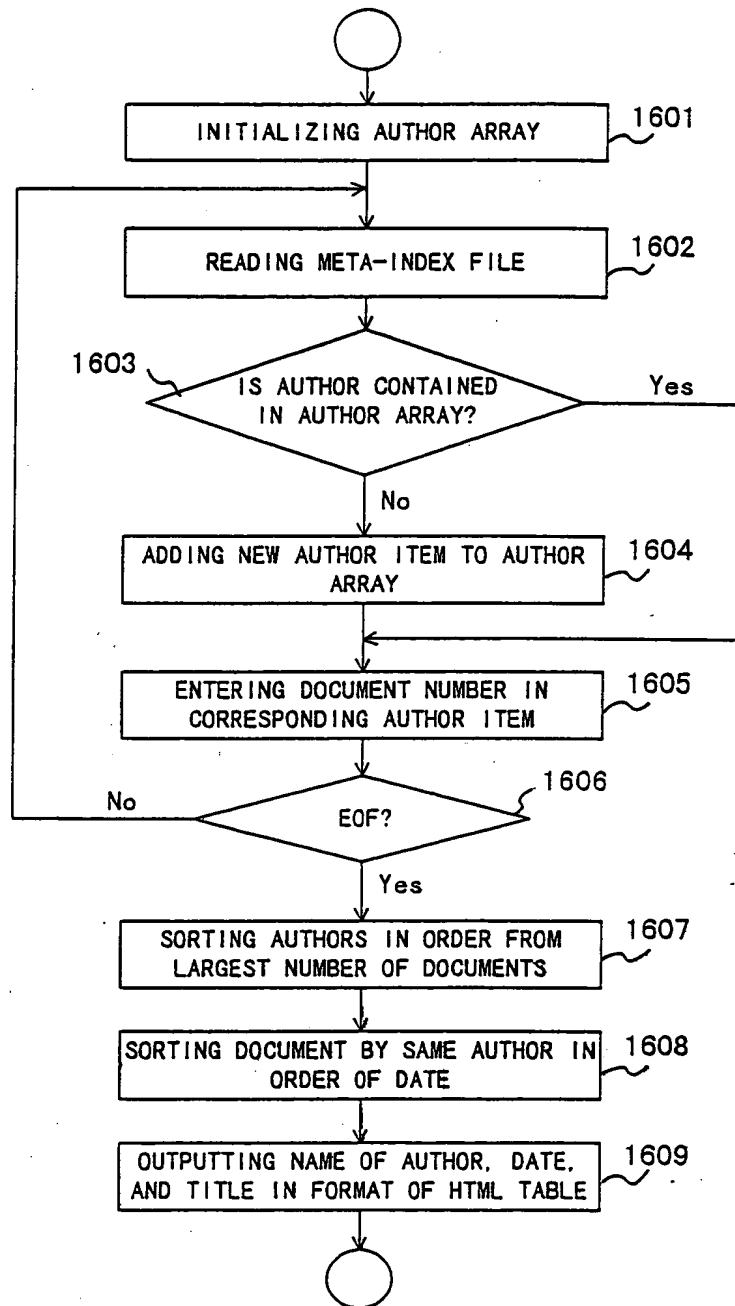
[FIG. 14] FLOWCHART SHOWING THE OPERATION OF A THREAD VIEW

LIST IN TREE STRUCTURE: (001 (002 003) (004 (005 006)))

	1	2	3	4
1	001	002	003	BLANK
2	BLANK	004	005	006

ARRAY INDICATING TREE STRUCTURE

[FIG. 15] EXAMPLE OF A CONTROL ARRAY FOR A THREAD VIEW

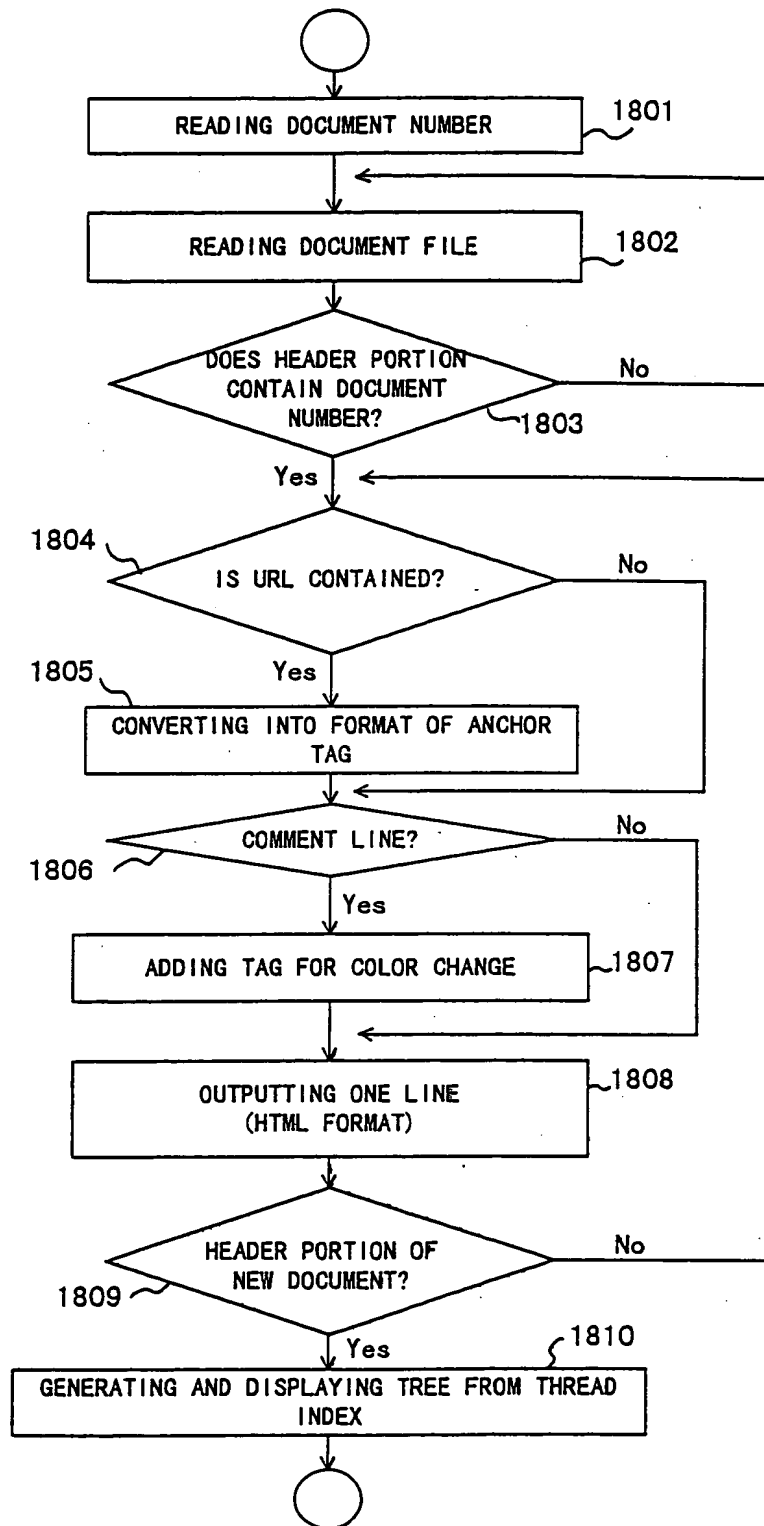


[FIG. 16] FLOWCHART SHOWING THE OPERATIONS OF AN AUTHOR VIEW

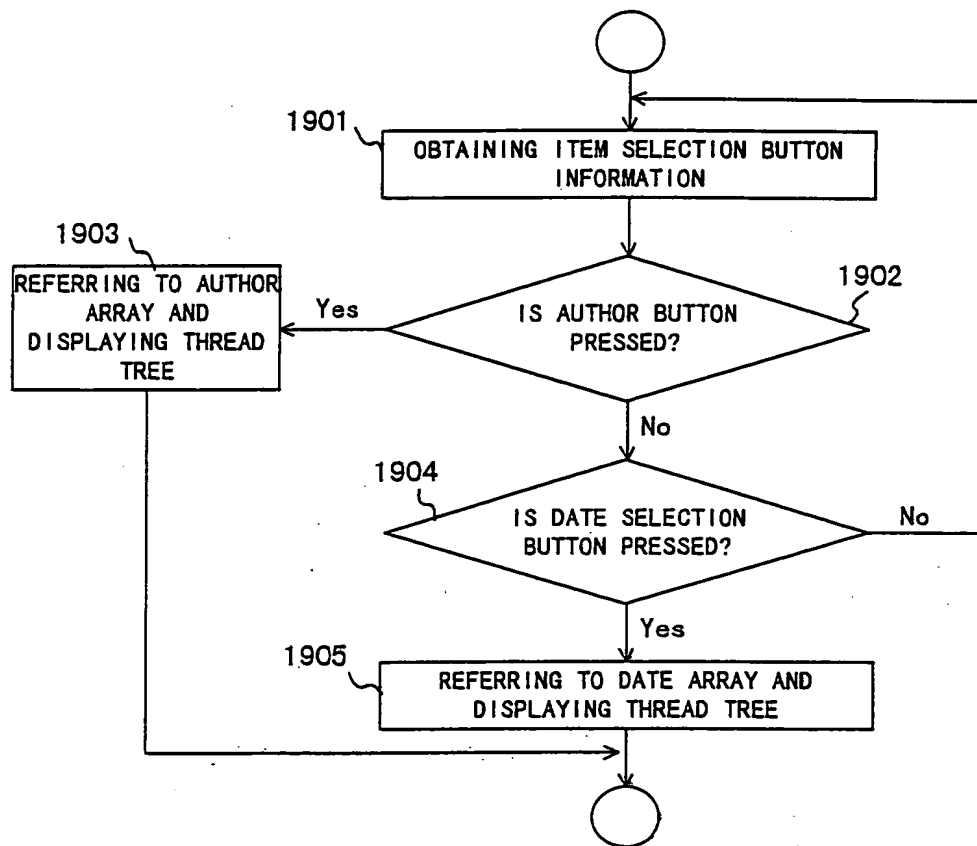
EXAMPLE OF AUTHOR ARRAY

AUTHOR A	001, 005, 007, 008
AUTHOR B	011, 012, 013
AUTHOR C	003, 009, 010
AUTHOR D	004, 006
AUTHOR E	002

[FIG. 17] EXAMPLE OF A CONTROL ARRAY FOR AN AUTHOR VIEW



[FIG. 18] FLOWCHART SHOWING THE OPERATIONS FOR
DISPLAYING THE CONTENTS OF A MESSAGE



[FIG. 19] FLOWCHART SHOWING THE OPERATIONS OF DISPLAYING AUTHORS AND DATES IN RESPECTIVE COLORS

EXAMPLE OF AUTHOR ARRAY

AUTHOR A	001, 005, 007, 008	→ YELLOW
AUTHOR B	011, 012, 013	→ BLACK
AUTHOR C	003, 009, 010	→ RED
AUTHOR D	004, 006	→ BLUE
AUTHOR E	002	→ GREEN

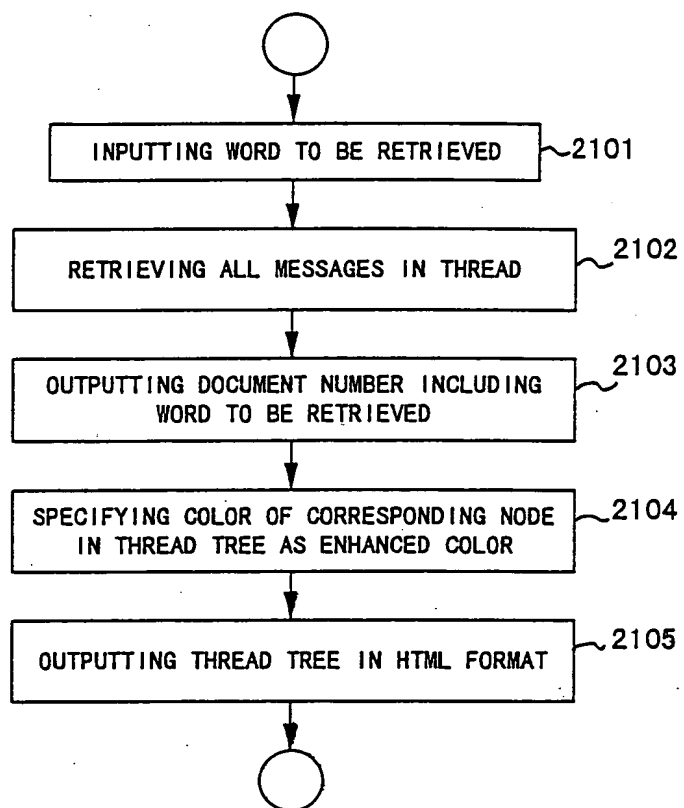
EXAMPLE OF DATE ARRAY

97/04/15	001, 002, 003, 004	→ YELLOW
97/04/21	005, 006	→ BLACK
97/04/28	007, 008, 009	→ RED
97/05/25	010	→ BLUE
97/05/26	011, 012, 013, 014, 015	→ GREEN

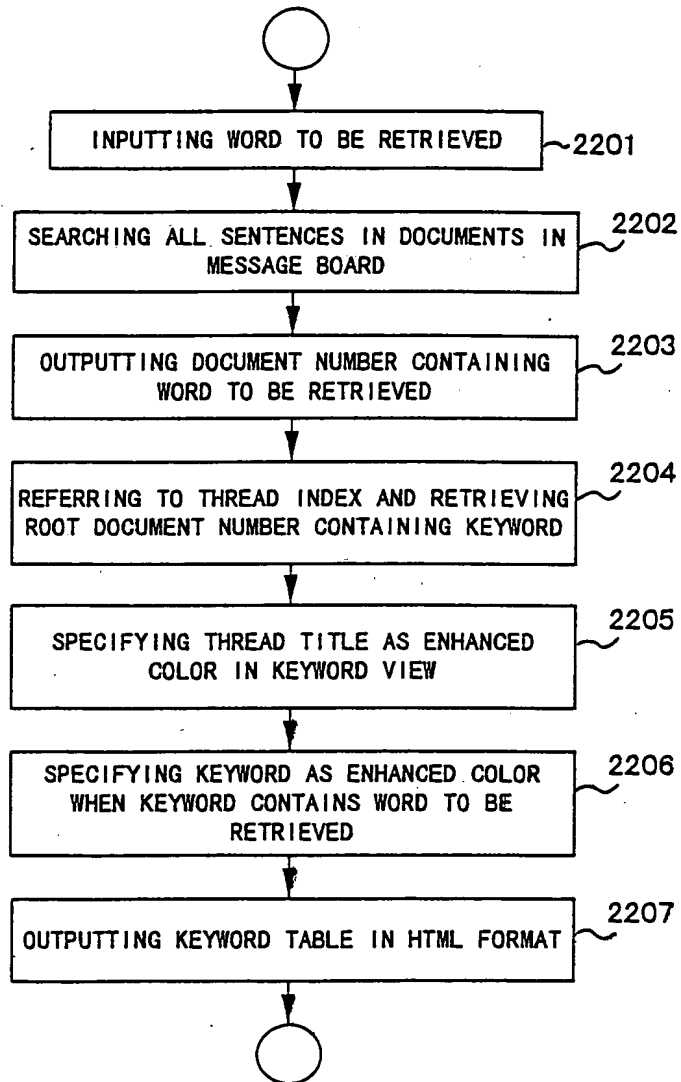
FIG. 20A

FIG. 20B

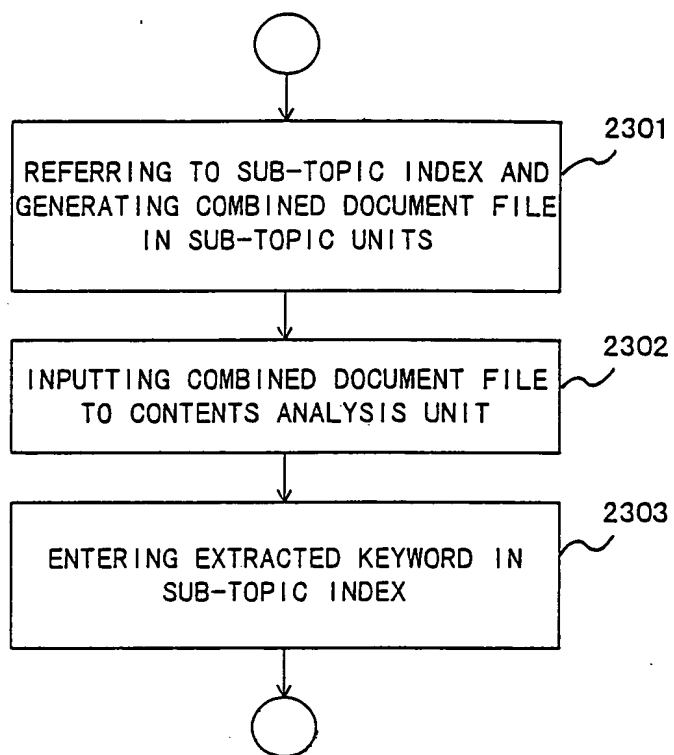
[FIG. 20] EXAMPLES OF DISPLAY ARRAYS FOR AUTHORS AND
DATES IN RESPECTIVE COLORS



[FIG. 21] FLOWCHART SHOWING THE OPERATIONS OF A HIGHLIGHTED DISPLAY OF A RETRIEVAL RESULT USING THREAD VIEWS



[FIG. 22] FLOWCHART SHOWING THE OPERATIONS OF A HIGHLIGHTED DISPLAY OF A RETRIEVAL RESULT USING KEYWORD VIEWS



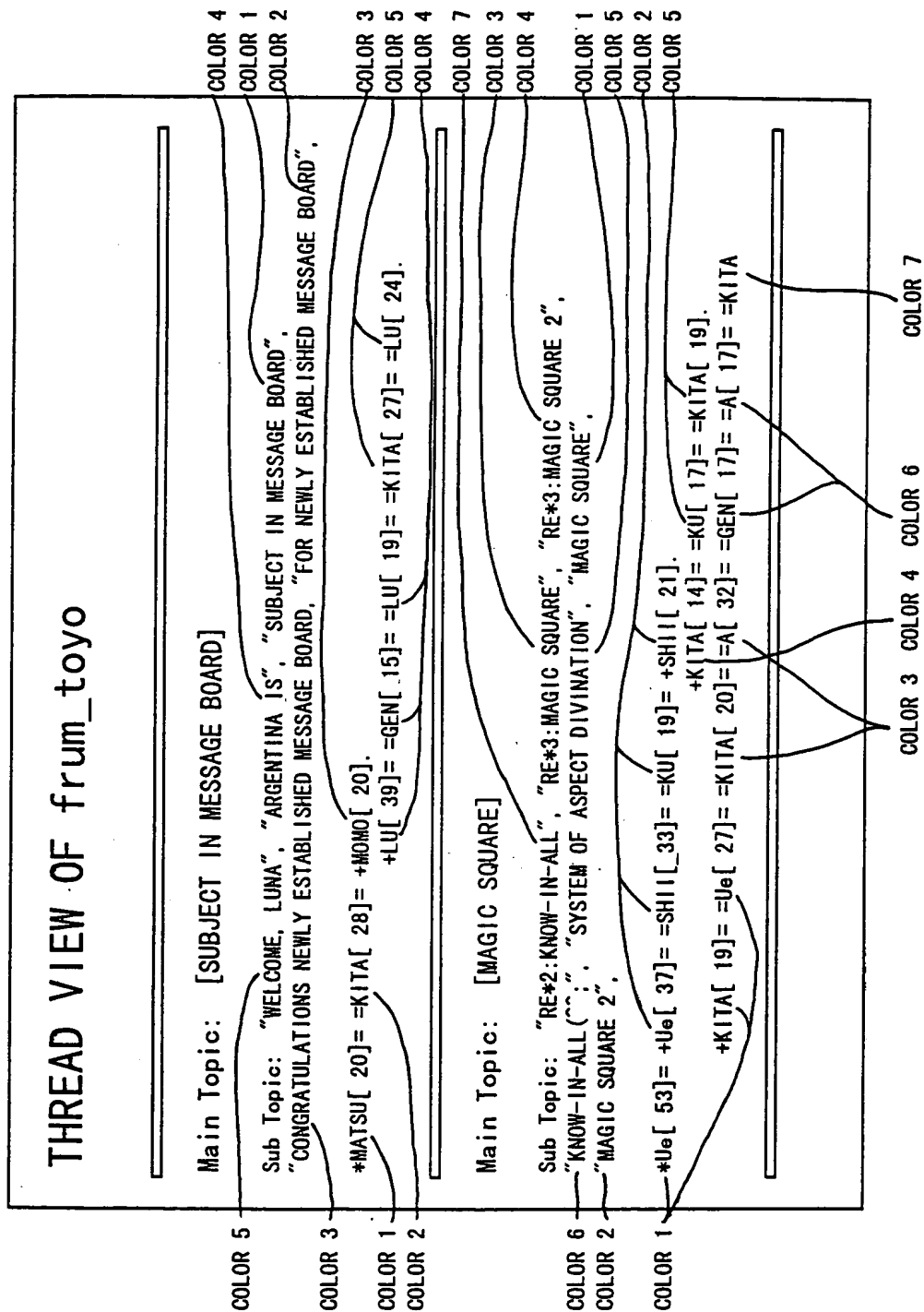
[FIG. 23] FLOWCHART SHOWING THE OPERATIONS OF CONTROLLING THE EXTRACTION OF A KEYWORD FROM A SUB-TOPIC

ROOT DOCUMENT NUMBER	COLOR NUMBER	TITLE	MESSAGE NUMBER	KEYWORD
:	:	:	:	:
001	#1	(MAIN TOPIC)	001, 002, 004
	#2	(SUB-TOPIC 1)	003
	#3	(SUB-TOPIC 2)	005, 006
:	:	:	:	:

[FIG. 24] EXAMPLE OF A SUB-TOPIC INDEX

KEYWORD VIEW OF fgal_internet_QA			
[THREAD VIEW]		[SENDER VIEW]	
ID		TITLE	KEYWORD
00001	<input type="checkbox"/> 1	PREFACE	CONTENTS OF QUESTIONS / THIS MESSAGE BOARD / TITLE OF MESSAGE / Internet / RELEVANT QUESTIONS AND ANSWERS / CONNECTION TYPE / ENVIRONMENT FOR ... (HARDWARE / APPROPRIATE ADVICE FOR CONTENTS / BABA / INABA / Manager
00002	<input type="checkbox"/> 3	MINIMAL HARDWARE	NET SURFING LATELY IN FASHION / INTERNET / DEFACT STANDARD NETSCAPE NAVIGATOR / PENTIUM MACHINE / THIS WAY / PC COMMUNICATION IF ... / PROVISIONS OF LAW / VARIOUS SERVICES / Worldtalk OF People / MADI
00005	<input type="checkbox"/> 3	QUESTION: CANNOT BE CONNECTED TO HOMEPAGE	PROVIDER / CONFIRMATION OF PASSWORD / APPLICABLE MODEL / SETTINGS / DEFAULT OF NETSCAPE / DIAL UP IP CONNECTION / APPLICABLE SOFTWARE / FMV*DPSP5 / USER NAME
00008	<input type="checkbox"/> 6	INTERNET START	PROVIDER / HOME PAGE / FEE / STARTING TIME OF TELEHODAI / WWW / ACCESS / ISDN / ANALOG 288 / BAUD RATE / nifty COMMAND
00014	<input type="checkbox"/> 1	THANKS EVERYBODY	PROVIDER RECOMMENDED BY REYA / IMPOSSIBLE NETWORKER / CALL TOMORROW / DISCOUNT OF GOLD FOR EMPLOYEES / SUMMARIZED ... / CHALLENGING IN MANY FIELDS / ... IN NAGOYA / ROAD55 / OTHER DAY / SUBSIDIARY
00015	<input type="checkbox"/> 24	NIFTY DIAL UP CONNECTION	@niftyserve / ISN'T IT? / SCRIPT / SubManager / or / PEOPLE FROM VENUS / NIFTYServe / INTERNET / WIZARD / kazusi
00016	<input type="checkbox"/> 1	SELECTING PROVIDER	IMPOSSIBLE NETWORKER / ONLINE SI... / RECOMMENDED BY REYA / SETTING HOME PAGE / REQUEST FOR MATERIAL / ... UP / FIXED FEE / EXTENT RELIABLE / RATHER HARD / ALLESNET
00030	<input type="checkbox"/> 9	PROVIDER ... BAUD RATE BETWEEN HOME PC	SERVER-SERVER HOME PC / @niftyserve / PROXI... / T*K / ping / MNP5 / CACHE FOR LOWER LAYER / SubManager / PEOPLE FROM VENUS / MODEM COMMAND
00040	<input type="checkbox"/> 24	NIF CONNECTION MAIL	INTERNET MAIL/ @niftyserve / SubManager / or / BABA / INABA / PEOPLE FROM VENUS / NIF-CONNECTED MEMBER / NIFTYServe / IN SERVICE / kazusi

[FIG. 25] EXAMPLE OF DISPLAYING A KEYWORD VIEW



[FIG. 26] EXAMPLE OF DISPLAYING A THREAD VIEW

		8: [RE^4: DOES ANYBODY KNOW SANYNET?]
MURA	SEPTEMBER	11: [Admin AUTHORIZATION FOR RE>NT] 30: [QuickTime WITH NESUKE3.0iβ6]
	OCTOBER	1: [RE^2: QuickTime WITH NESUKE 3.0iβ6] [RE>NN3Jaβ CANNOT BE INSTALLED] 2: [RE>QuickTime WITH NESUKE 3.0iβ6] [RE>QuickTime WITH NESUKE 3.0iβ6]
KANCHAN	JULY	25: [WANT TO RED news WITH 98+DOS] 27: [WANT TO RED news WITH RE: 98+DOS]
	SEPTEMBER	29: [ABOUT EXTENSION rar] [RE: MAIL WITH NO SENDER NAME] 30 [RE: ABOUT EXTENSION rar]
	OCTOBER	1 [RE: ABOUT EXTENSION rar]
YOSHI CHAN	MAY	6: [RE^1: NIFTY DIAL UP CONNECTION] 18: [RE^5: NIFTY DIAL UP CONNECTION]
	AUGUST	5: [RE^2: DOES ANYBODY KNOW SANYNET?] 6: [RE^3: DOES ANYBODY KNOW SANYNET?]
OSAMU ARAKI	AUGUST	3: [PRIMARY QUESTION ABOUT ActiveX] 8: [RE: PRIMARY QUESTION ABOUT ActiveX] 31: [IE3 FORMAL VERSION TROUBLE]
	SEPTEMBER	2: [RE: IE3 FORMAL VERSION TROUBLE] 3: [RE: IE3 FORMAL VERSION TROUBLE]
TOSHIHIRO SAKAWA	SEPTEMBER	9: [CANNOT BE CONNECTED TO INTERNET (2)] 14: [CANNOT BE CONNECTED TO INTERNET (2)] 15: [RE:RE^2: CANNOT BE CONNECTED TO INTERNET (2)] 16: [RE:COMMON CONNECTION METHOD]
Mue	AUGUST	4: [RE: QUESTION HOW CAN HOME PAGE BE PRINTED ...] 18: [RE: PPP COMPARISON BETWEEN Hyper Road AND Road4] 25: [RE: ABOUT INET SETUP WIZARD]
SADAYUKI ANDO	JULY	27: [FELL IN TROUBLE SOS!] 28: [RE: FELL IN TROUBLE SOS!] 30: [RE: FELL IN TROUBLE SOS!]
N-DATE	AUGUST	18: [ABOUT RealAudio PLUG IN] 19: [RE: ABOUT RealAudio PLUG IN] 25: [RE: ABOUT INSTALLING RealAudio]

[FIG. 27] EXAMPLE OF DISPLAYING AN AUTHOR VIEW

PROVIDER - - - BAUD RATE BETWEEN HOME PC

00030/00524 AAA00001 AAA
(12) 96/05/10 03:02

PROVIDER - - - BAUD RATE BETWEEN HOME PC
NUMBER OF COMMENTS: 1

WHEN WE SEE Homepage, SERVER-SERVER-HOME PC CONNECTION IS TO BE MADE?
(ACTUALLY THERE MUST BE OTHER METHODS)

WHAT'S THE BAUD RATE BETWEEN SERVERS? IS IT 14400bps (IN MY CASE) SET
USING MODEM COMMAND?

SINCE DATA ARE OVERFLOWING IN MANY CASES ...

AAA (AAA00001)

THREAD

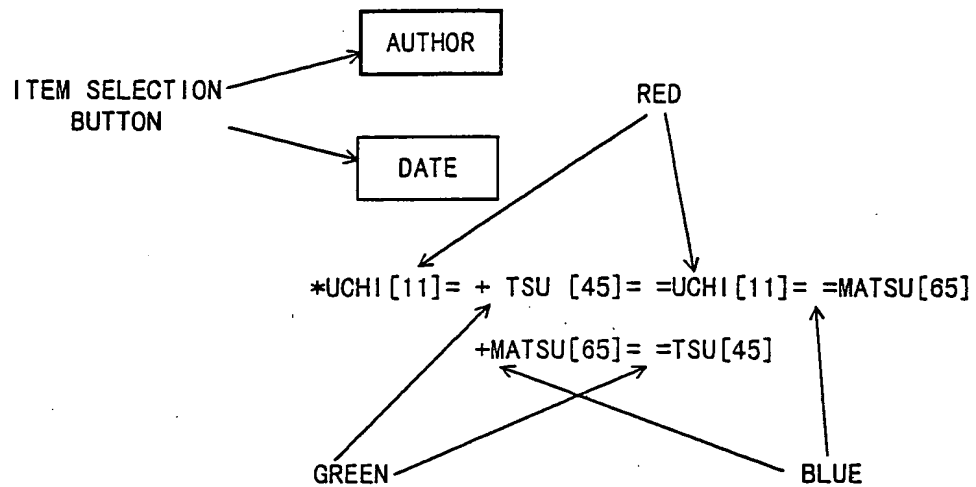
*00030= 00031= 00034= 00035= +00038= 00054.
+00066= 00067= 00068.

<<Prev---thread---Next>>

[KEYWORD VIEW] [THREAD VIEW] [SENDER VIEW]

RETRIEVAL FROM ENTIRE
MESSAGE BOARD

[FIG. 28] EXAMPLE OF DISPLAYING THE CONTENTS OF A MESSAGE



[FIG. 29] EXAMPLE OF DISPLAYING WITH HIGHLIGHT (IN DIFFERENT COLORS) A DOCUMENT ATTRIBUTE 'AUTHOR' USING A THREAD VIEW

DOCUMENT RETRIEVAL PAGE IN fgal_internet_QA MESSAGE BOARD

THIS SERVER SEARCHES DOCUMENT RESOURCES OF fgal_internet_QA MESSAGE BOARD.
ENTER RETRIEVAL CHARACTER STRING AND RETRIEVAL CONDITION IN THE FOLLOWING FORMAT,
AND CLICK THE SEARCH BUTTON.

SINCE DATA CAN BE SELECTED FROM A RETRIEVAL RESULT, IT IS RECOMMENDABLE NOT TO
START RETRIEVAL UNDER SERVER CONDITIONS.

DOCUMENT CONTAINING ANY OF

EXCLUDING

..... IS TO BE RETRIEVED.

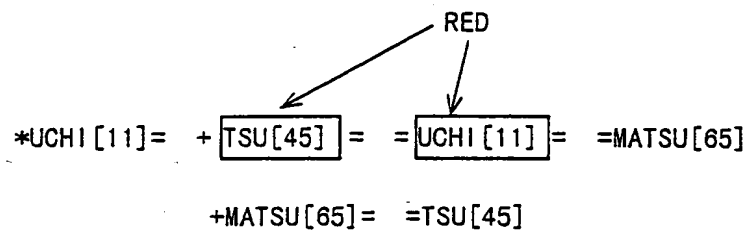
THE UPPER OR LOWER CASE OF ALPHABETICAL CHARACTERS SHOULD/SHOULD NOT BE
DISTINGUISHED, AND BLANK CHARACTER STRING IS RETRIEVED AS IGNORED /
WITHOUT IGNORED.

[RETURN TO FIRST PAGE](#)

[ABOUT....](#)

[FIG. 30] EXAMPLE OF DISPLAYING RETRIEVAL (INPUT) OF
ALL MESSAGES IN A MESSAGE BOARD

WORD TO BE RETRIEVED: PROTOCOL



[FIG. 31] EXAMPLE OF DISPLAYING WITH HIGHLIGHT A NODE CONTAINING A CHARACTER STRING 'PROTOCOL' USING A THREAD VIEW

WORD TO BE RETRIEVED:PROTOCOL	
NAME OF MESSAGE BOARD:FGL INTERNET	
TITLE	KEYWORD
A OUR TEACHER ... CABLE	B UNIVERSITY C SCHOOL BUILDING START OF INTERNET
PROTOCOL	telnet ftp PROTOCOL OF UNIX MACHINE
...	...
STANDARD OF Internet	TYPE OF PROTOCOL http ftp
...	...

RED

[FIG. 32] EXAMPLE OF DISPLAYING WITH HIGHLIGHT A THREAD CONTAINING A CHARACTER STRING 'PROTOCOL' USING A KEYWORD VIEW

Result: FOUND

IN THE DOCUMENT RESOURCES OF THE fgal_internet_QA MESSAGE BOARD, THERE ARE 14/518 ITEMS SATISFYING LOGIC EXPRESSION (UNIX) 8 NETSCAPE & INTERNET & HOMEPAGE.

1. 00064 : RE^2: PROTOCOL
2. 00091 : RE: chmod IS
3. 00204 : RE: TELL ME HOW TO DOWNLOAD NIF.
4. 00212 : RE^2: TELL ME HOW TO DOWNLOAD NIF.
5. 00233 : RE: RE: I HAVE GIVEN YOU INFORMATION ...
6. 00270 : RE^2:
7. 00308 : TIY
8. 00330 : RE: RE^2: TELL ME HOW TO DOWNLOAD NIF.
9. 00343 : BPULS FOR UNIX
10. 00349 : RE: PREVENTING PROCESS FROM BEING LEFT OVER
11. 00376 : RE: MONITOR CANNOT BE SEEN?
12. 00463 : RE: Q HOW CAN EXTENSION WITH taz
13. 00475 : RE^2: HOW CAN EXTENSION WITH taz
14. 00480 : RE^4: Q HOW CAN EXTENSION WITH taz

SELECTIVE RETRIEVAL

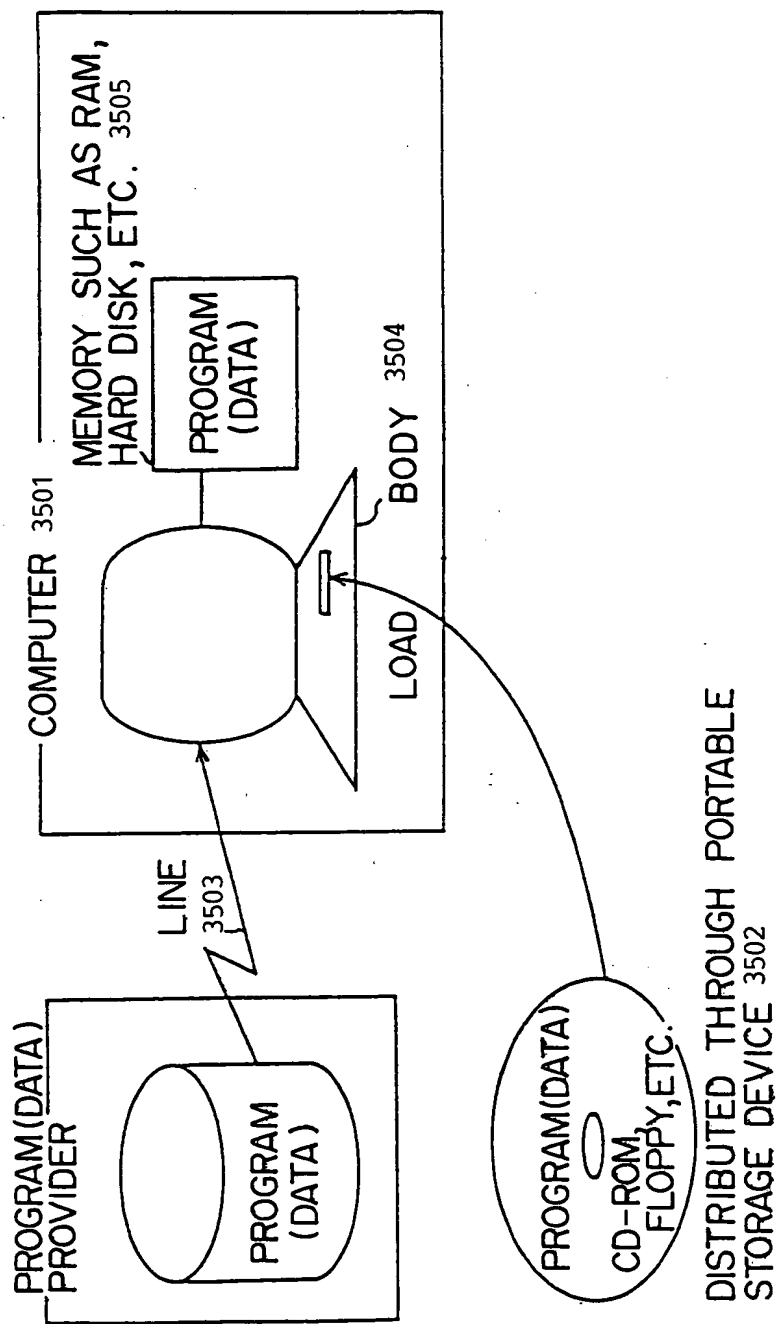
SELECTIVE RETRIEVAL CAN BE MADE BY SETTING CONDITIONS FROM THE RETRIEVAL RESULTS. CLICK THE SEARCH BUTTON WITH THE FOLLOWING CONDITIONS SPECIFIED. WHEN NEW RETRIEVAL IS STARTED, RETURN TO THE FIRST PAGE.

[FIG. 33] EXAMPLE OF DISPLAYING RETRIEVAL (OUTPUT OF RESULT) OF ALL MESSAGES IN A MESSAGE BOARD

MAIN TOPIC: WHAT IS THE FIRST QUESTION?

SUB-TOPIC	KEYWORD
A OUR TEACHER ... CABLE	B UNIVERSITY C SCHOOL BUILDING START OF INTERNET
PROTOCOL	telnet ftp PROTOCOL OF UNIX MACHINE
.....

[FIG. 34] EXAMPLE OF DISPLAYING A KEYWORD EXTRACTED FROM
A SUB-TOPIC



[FIG. 35] STORAGE MEDIUM STORING A PROGRAM FOR REALIZING
A PREFERRED EMBODIMENT OF THE PRESENT INVENTION

[Document Name] Abstract

[Abstract]

[Subject] The present invention aims at supporting the retrieval of required documents or a group of documents
5 probably containing required documents from a set of documents containing a great amount of documents.

[Means for solving the Problems] A group of documents in a forum/message board are stored as a document group database 301 in a predetermined server computer. A
10 document group analysis device 302 analyzes a group of documents contained in each set of documents corresponding to each message board in the document group database 301. A summation device 303 generates a meta-index 304, a thread index 305, and an indexed
15 file 404 based on analyzed results of the document group analysis device 302. A display device 306 displays a group of documents in a display format of a keyword view, a thread view, or an author view using a meta-index 304 and a thread index 305.

20 [Selected Drawing] FIG. 3